

FIGURE 1

GGGGCTCGGCCAGCGCCAGCGCTAGTCGGCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCCTCAGTTCTTCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTATGCATTG
CTACCATTTATGTTGTTATAAGCAAGTCATGCTCGAGTCAGTGCAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGGCAAACCTT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTAGACCATCCTTCTACCAATGCAGCCCCAAATCCATGGCAAACAGTC
TTCTGGATCAGACTGTTGGTTATCTGGTGTGGAGTAAGTGCACTTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAAACCTCCATTGGAACCCGGAGG
ACAAAAGGTTATGTGCTTCACATGACTACTGCAGCAGAACGGCTATGTCAATTCTCTTT
GGTTTTCTGACTTACATCGTATTTCAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTGCCCTATTAACAATGAACGAACACGGCTACTTCCA
GAGATTTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAACACCTATGCCCTACCTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCCCGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCAAGCGTGTGGCCTGGGGAGTGGAGTGGAGGCAGGAGCCTTC
CTTACACCTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTTCTCATGCCAATTGTTAAAGACTATGAGATACGTAGTATGTTGTACAG
GTGATCTCTCCGTGACGTTGCATTTCCTGCACCATGTTGAGCTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTCATGGTGCCTTTTACATTGGCTATTGAGCAATATCCGACTACTGCATAAACACGA
CTGCTTTCTGTCTCTTATGGTGCACCTTATGTATTCTTGAAACTAGGAGATCCCTTCC
CATTCTCAGCCCCAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTCTTGATTTGGCTGTCAACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGGATATTCTAGCCCTGGAACGGCAGCTGCTGCAAACCATGGATATGAT
CATAAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAGAAATTAAAGCAGGCAGCTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAAATAGAATACTCCAAACCTTCAAGGGAAATATTTAATTTCCTGGTT
ACTTTCTCATTTACTGTGTTGGAAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCAACACATTCTCATTCTGTTGAATAATCATGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTCTTTATGCCATCTCTAGCAGTAAGTCTCCAATGTCATTGCTCTG
CTATTAGCACAGATAATGGGATGTACTTTGTCCTCTGTGCTGATCCGAATGAGTATGCCCTT
AGAATACCGCACCATATAACTGAAGTCCTGGAGAACACTGCAGTTCAACTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCCTGACATTATAAAACAAACAAATGCTATGGTAGC
ATTTCACCTTCATAGCATACTCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCCAAGAAACTAAAGGTGAAAATACACTGGAACTCTGGGCAAGACATGT
CTATGGTAGCTGAGCCAACACGTAGGATTCCGTTTAAGGTTACATGGAAAAGGTTAGCTT
CCTTGAGATTGACTCATAAAATCAGAGACTGTAACAAAAAAAGGGCGCCGCG
ACTCTAGAGTCGACCTGCAGAGCTTGGCCCATGGCCAACTTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWMNLCVILLILVFMPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFFWKLGD
FPILSPKHGILSIEQLISRVGVIGVTMALLSGFGAVNCPTYMSYFLRNVTDTDILALEERRLLQ
TMDMIISKKRMMARRTMFQKGEVHNKPSGFWMKSVTTSASGSENLTLLIQQEVDALEELSRO
LFLETADLYATKERIEYSKTFKGKYFNFGLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQMIMGMY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCACTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCAGT
TCCATCTGGACCACGAGGCTCTGGTCCAAGGCTTTGCGTGAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGATCACCCCTGTGAGCAAAAAGGCGAACAGCAGCTGAATTACAG
AAGCTAAGGAGGCTGTTAGGCTGCTGGACTAAGTTGGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAAACTTGCAGCTATGGCTGGGTTGGAGATGGATTCTGTGGTCATCTCTAG
GATTAGCCAAACCCCAAGTGTGGAAAAATGGGTTGGTGTCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCAGCTATTGTTACAACACTCATCTGATACTTGGACTAACTCTGTCATTCCAGAA
ATTATCACCAACAAAGATCCATATTCAACACTCAAACTGCAACACAAACAGAATTATTGT
CAGTGACAGTACCTACTCGTGGCATTCCCTACTCTACAATACCTGCCCTACTACTACTCC
CTGCTCCAGCTTCACTCTATTCCACGGAGAAAAAAATTGATTGTGTGTCACAGAAGTTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAAACCATTTGTTGAAATAAAAGCAGCATTCAAGAATGA
AGCTGCTGGTTGGAGGTGTCACGGCTCTGCTAGTGCCTGCTCTCTCTTGGTGTG
CAGCTGGCTTGGATTGGCTATGTCAAAGGTATGTGAAGGCCCTCCCTTACAAACAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCATGATAGCAACCTAA
TGAGGAATCAAAGAAACTGATAAAACCCAGAAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGC
GATGCCCTGGAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGGCCAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAACAGTCCA
CCCTGGTTCTTAACCTGGAAATCAGCTCAGGACTGCCATTGGACTATGGAGTCACCAAGAGAAT
GCCCTTCTCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCACGGCCT
TTCTAGCCTGGCTATGCTCTAATAATATCCCACGGGAGAAAGGAGTTTGCAAGTGCAGGAC
CTAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCCTGGCTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCAGTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTCTGAGCCCGTAAGAGCAAAGAAT
GGCAGAAAAGTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTGTAAA
GCTAAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTAGCAGAGGACTGTAAC
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGGAATCACTGTTAGAACACACACA
CTTACTTTCTGGTCTCTACCAACTGCTGATATTTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAACCTTATAAATTCTATTCTAGGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCTCTAAAGGAAATTGCA
AACGCTATCTGGAGCTATTCTGAGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTCTGAGACTAATCTTATTCTATTCTCTAATATGGCAACCATTATAACCTTAATT
TATTATAACATACCTAACAGAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGTGC
ATTAAACAAATGTATCACTAGCCCTCCTTTCCAAACAAGAAGGACTGAGAGATGCAGAAATT
TGTGACAAAAATTAAAGCATTTAGAAACTT

FIGURE 6

MARCFSLVLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNSPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTEPEFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCCGCCTCCGCACCCGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGCAGCCCGGC
GGCCTCCCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGTCCAGTCGGGGCG
CGGCTCGGGCGCAGAGCGAGATGCAGCGGTTGGGCCACCTGCTGTGCCCTGCTGGCG
CGCGGTCCCCACGGCCCCCGCGCCGCTCGACGGCACCTGGCTCAGTCAGGCCCG
GCTCTAGCTACCCGAGGAGGAGGACCCCTAATGAGATGTCCGAGGTTAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGGTGGAAGAGATGGAGGAGAAAGACTGCTGTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTAACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGACCGAGAAATCACAAGATAACCAACAAACAGAC
TGGACAAATGGCTTTCAAGAGACAGTTACACATCTGTGGAGACAGAAGGGAGAAGGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTGCCAGCTTCCAGTAC
ACCTGCCAGCATGCCGGCCAGGGATGCTCTGCACCCGGACAGTGAAGTGTGTGGAGACCA
GCTGTGTCTGGGTCACTGCACCAAAATGGCCACCCAGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCGGGCTGTCTGTGCTTCCAGAGAGGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCGTGGAGGGCAGCTTGCCATGACCCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGACGCTGATGGAGCCTGGACGATGCCCTGTGCCAGTGGCCTCTGCCAGC
CCCACGCCACAGCCTGGTATGTGTGCAAGCCGACCTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCGATGAGTATGAAGTGGCAGCTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCGTGCG
CCGCCGCTGCACTGCTGGGGAGGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGAA
TAGAAATAGCTAATTATCCCCAGGTGTGCTTAAAGCTGGCTGGCTGACCAAGGCTCTCCTA
CATCTTCTTCCAGTAAGTTCCCCCTGCTGCTGACAGCATGAGGTGTTGCAATTGTCAGCT
CCCCCAGGCTGTTCTCCAGGCTTACACTGTTGCTTGGAGACTGAGCTGGCAGACAGCCG
GGAGCAGTTGCCACCCCTGTCAGATTATGGCTCTGCTTACCTACAGTGGCAGACAGCCG
TTTGTCTACATGGCTTGTATAATTGTTGAGGGAGGAGATGAAACAAATGTGGAGTCTCCCTC
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGTTGCAACACATCAACCTGGCAAAATG
CAACAAATGAATTTCACCGCAGTTCTTCCATGGCATAGGTAAGCTGTGCCCTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCAATTACATGTGTTATTCACTCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCAGGGCAGCAATTCTCATATCCAAGATCAATTCCCTCTCAGCACAGCCTGGGG
AGGGGGTATTGTTCTCCTGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTACTCTAAGCTCAGTGTCT
CTCCACTACCCACACCAGCTTGGTGCCACCAAAAGTGCCTCCCCAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGGAAATTAGTCAGTAACTAATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGCATTAGTAACCTTGAAAGGTATATGACTGAGCCTAGCA
ATTGACACTGTCCTCTTGGCAGTTGCAATTGAGGGCAGGATTAAATGAAATTGCA
TACAGGTTAACCTGCAGAACAGTACTTAGTAATTGAGGGCAGGATTAAATGAAATTGCA
AAAATCACTTAGCAGCACTGAAGACAATTATCAACCCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATTGCGACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTCTGACTGGACTGTTGCCACCATGTATTCACTCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAAACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDKVGNNTIHKITNNQTGQMFSE
TVITSGDEEGRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRLMLCTRDECCGDQLCVWGHCT
TKMATRGSNGTICDNQRDCQPGGLCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGCTGTGAGAAAGTCCAATAAATACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGTGCCTCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTTCTGGACCCCTT
AACTGGGTACTGCCCTGGCCAATGCGTCCTCGCTGGAGGCTTGCCTCCTCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCTTAATCTCTGCCTCATCCGCACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGGTATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCTGTAGCCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTCTAAACCGCAATGCATACTCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCCTCATGCTACTCATGCGAAC
ATTGTCAAGGGTGGTCGTCCCTGGACAAAGTCACAGACCTGCTGTTCTTGGGAAGCTGCTGGT
GGTGGAGGCGTGGGGTCTGTCTTCTTCTCCGGTGCATCCGGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCCTAT
GTCATGCCAGCGGCTTCTCAGCTTTCGGCATGTGTGGACACGCTCTCCTCTGCTTCC
GGAAGACCTGGAGCGAACACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCCGAACACAAGAAGAGGAAGAAGT**GACAGCTCCGG**
CCCTGATCCAGGACTGCACCCCACCCACCGTCCAGCCATCCAACCTCACTCGCCTTACAGGT
CTCCATTGTGGTAAAAAAAGGTTTAGGCCAGGCGCGTGGCTACGCGTGTAAATCAAACACT
TTGAGAGGCTGAGGCGGGGGATCACCTGAGTCAGGAGTTCGAGGACAGCAGCTGGCCAACATGGTG
AAACCTCCGTCTATTAAAAATACAAAAATTAGCCAGAGTGGTGGCATGCACCTGTATCCCA
GCTACTCGGGAGGCTGAGGAGGAATCGCTGAACCCGGAGGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTATTAAGATATTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLRNAYIMIAIYGKNFCVSAKNAFMLMRN
IVRVVVLDKVTDLLFFGKLLVVGGVGVLSSFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLKILGKKNEAPPDNKKRK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGGCCGGCGCCGGCGCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTCACGTTCTTCCTGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGCCTGCAGGCCACATCGACTGTGGCTCCCTGCTGGTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGCCTCTCTTCTTACCCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGCTGCCATCCAGAATGGGTTTGTTCTTAAGTTCTGATCCTGGTGGCCTC
ACCGTGGGTGCCTCTACATCCCTGACGGCTCCTCACCAACATCTGGTTCTACTTCGGCGTGTGG
CTCCTTCCTCTTCATCCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTCTGGAAACCAGCGG
GGCTGGGCAAGGGCAGGGAGTGCATTCCCGTGCCTGGTACGCAGGCCCTTCTTCACTCCCTC
TTCTACTTGCTGTCGATCGCGCCGTGGCGCTGATGTTCATGTAACACTGAGCCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTCTGTGTCTGCGTGTCCATCGCTGCTGTC
CCAAGGTCCAGGACGCCAGCCAACCTGGGTCTGCTGCAGGCCCTGGTACATCACCCCTACACCATG
TTTGTCAACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCTCATCATCTCCTCTGTGCACCTCTTCATCAGCTGCGCTCTCAGACCACGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCACCTATGCTAGACGCCACACGAGCAGCAGCAGCAGG
GGCAGCCTGTGAGGGCCGGCTTGACAACGAGCAGGACGGCGTACCTACAGCTACTCCTCTTCC
ACTTCTGCCCTGGTGTGGCTCACTGCACGTATGATGACGCTACCAACTGGTACAAGCCCCGTGAG
ACCCGGAAGATGATCAGCACGGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCGCAACCGCGACTTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTGGCTCGGTGACGCCAACCTGCC
CCCCACCCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGGCC
CTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCCTTCCTCCCTCCTGGTGCCTGAGGAGCATCAGCAGCAGCAG
AGGGCTCCCTGTCCCTCAGGCTCCAGGGAGGGGCTGCTGGAGAGAGGCCACTCCACACAG
TGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTACGTCCCCAGGGACCCCTGCCCTTCTG
GACTTCGTGCCTTACTGAGTCTTAAGACTTTCTAATAACAAAGCCAGTGCCTGAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLIMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHAWNQRWLGKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCFLFISLRSSDHQRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGTVYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICASWAGLLLILWTLVAPLLRNDF
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGCGGCCAGGAACCACCCGTTAAGGTGTCTCTCTTACGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACCTGCCAGAACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTCTGTTGTTGTCACCTTGCACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAACGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTGATATATTCTCTGGAGTTTCGATTTAAAGTGTAAACTTGCATATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCTCAAGGGCTTTGGCTATGTGCTGCCATCATTCAATC
CTTGCCTGGATTGAGACGTGGCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTCAGAGAGGGCAGCACTTACCTGGTGGCTTGATG
GTCAGTTTATTCCCCCTCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTAGAACATATGACTACTTTGTTAAATGTGAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGTTTATTGAACAGCTAATAAGATTTATTGTAATACCTCACAAACGTGTAC
CATATCCATGCACATTTAGTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAACGCTTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGAACCTTTTATTCAATTCA
AGAATGGAATTTTTGTTCACTCATGCACATGTGCTTTGTCAGTTAAAAGTGTAAATAAAATCTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTGTATGGCTGAAGTGTGGA
CTTGCAAAAGGGAAAGAAAGGAATTGCGAACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGATGC
ACAAAATGACTAAACCATTCAATCATGTTCTGAAATGTTCTAAATACTCTTATTGATGC
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCGACCCGCCAGGAAAGACTGAGG
CCGGCCCTGCCCGCCGGCTCCCTGCAGCCGCCCTCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAGCCACAGACAGTCTTCTGCACTGCCGCCAGGGGACCACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGCTGTACGTCTTGAGAACGGCATCACCATGCTGACGC
AGGCAGCTTGCCGGCCTGCCGGCCTGCAGCTCTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGCGGGGCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATACCAATGAGACCTTCCGTGGCCTGCGGCCCTCGAGCGCCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTCGACAGCCTGACGCCCTCTGGAGCTCAAGCTGC
AGGACAACAGAGCTGCCGGACTGCCCGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCC
CACAACAGCCTCTGGCCCTGGAGCCGCATCTGGACACTGCCAACGTGGAGGCCTGCC
GGCTGGTCTGGGCTGCAGCAGCTGGAGGGGCTTCACTGCCCTGCGCAACCTCCACGACC
TGGATGTGTCGACAACAGCAGCTGGAGCGAGTGCACACTGCCCTGATCCGAGGCTCCGGGCTGACG
CGCCCTGCCGGCTGCCGGCAACACCCGATTGCCAGCTGCCCGAGGACCTGCCGGCCTGCC
TGCCCTGCAAGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCCTGCCACCTCTGGG
TCTTCCCCCGCCCTGCCGCTGCTGCCAGCTGCCCAACCCCTCACTGCCGTGCCCCCTGAGC
TGGTTTGGCCCTGGGTGCGCGAGAGCCACGTACACTGCCAGCCCTGAGGAGACGCCGTGCC
CTTCCCGCCAAGAACGCTGCCGGCTGCTCTGGAGCTTGAACGCCGACTTTGGCTGCCAG
CCACCAACACACAGCCACAGTGCCACACAGAGGCCCTGGTGCAGGCCACAGCCTTGCT
TCTAGCTTGGCTCTACCTGGCTTAGCCCCACAGGCCGGCCACTGAGGCCACGCCCT
CACTGCCCAACCGACTGTAGGGCTGTCCCCCAGCCAGTGCACGCCACCTGCCCT
ATGGGGCACATGCCACCTGGGACACGGCACCACTGGCGTGTGCCCCGAAGGCTTCACG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGACACGGCCAGCCCTACACCAGTCACGCC
GCCACACGGCTCTGACCTGGGATCGAGCCGGTGAAGCCACCTCCCTGCGCGTGGGCTGC
AGCGCTACCTCCAGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCACCTATCGCAACCTATCG
GGCCTGTATAAGCGGCTGGTACGCTGCAGCTGCCCTGCCCTGCCCTGCGTGAAGTACACGGTACCCA
GCTGCCGCCAACGCCACTTACTCCGTGTGTCATGCCCTGGGCCGGGCGGGTGCAGG
GCGAGGAGGCCTGCCGGAGGGCCATACACCCCCAGCCGTCCACTCCAACACGCCAGTCACC
CAGGCCCGCAGGGCAACCTGCCGCTCCATTGCCCGCCCTGCCCGGTGCTCTGGCCGC
GCTGGCTGCCGTGGGGCAGCTACTGTGCGGCCGGGGCCATGCCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGCCAGGGCTGGCCCTGGAAGTGGAGGGAGTGAAGGTCCCCCTGGAG
CCAGGCCGAAGGCAACAGAGGGCGTGGAGGCCCTGCCAGGCCGTCTGAGTGTGAGGTGCC
ACTCATGGCTTCCAGGGCTGGCTCCAGTCAACCCCTCCACGCCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGCCGGGCTCAGCCAGTGTGAGATGCCAGCCCCCTCCGTGCC
ACACCACTGTAAGTCTCAGTCCCAACCTGGGATGTGCAAGACAGGGCTGTGACCAACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCCTCATCTGTGAGATGCTGTGGCCAGCTGAGCC
CTAACGTCCCCAGAACGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGACCGGGGGCTGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
TCCAGGCCGGACCTGGGGCCAGTGAAGGAAGCTCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTACTCTAGTCTGGCCCCAGGAAGCAGAGGAACAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTGTCTTTAAATATATATTTATAAGAGATCCTTCCCATTATTCT
GGGAAGATGTTTCAAACCTCAGAGACAAGGACTTGGTTTGTAAAGACAAACGATGATATGAA
GGCTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLKLQDNELRALPPLRLPRLLLLLDSHNSLLALEPGILDANVE
ALRLAGLGLQQLDEGLFSRLRNHLHDVSDNQLERVPPVIRGLRGLTRRLLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLPGDLSGLFPRLLLAAARNPFNCVPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHACLCPEGFTGLYCESQMGQQGTRPSPTP
VTPRPPRSLTIGIEPVSPSTSRLVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASLA
TWTQLRPNATYSVCVMPGPGRVPEGEAACGEAHTPPAVHSNHAPVTQAREGNLPLLIAPALAA
LLAALAAVGAAYCVRRGRAMAAAQDKGQVGPAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCAGGCAGCGGTGGCTGAGTCGTGGCAGAGCGAAGGCACAGCTCATGCG
GGTCGGATAGGCTGACGCTGCTGTGCGGTGCTGAGCTTGGCTCGCGTCCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTGATTCAAAGAACTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACCGCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAAATGATGTACAAACTGAAATGAAAATCCTTAATGAAAG
CAATAAGAAAAGCCAAAAAAGAGAAGCATATCGGTATCTCAAAGGCAGCAAGCATGAAACCATA
CCAAAGCCCTGGAGAGAGTGTCAATGCTCTTTATTGGTGAATTGCCCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTCTGTATATTACAT
TTGGAGCTTGGGGCAACTAAATAGCCCACATGGTTTGGTAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAAACATCAGAAGAATTGTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTGTATTAAGACCAAATATTCAAGTCAGTGAACCTCCCTCAAATTCTGTAAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTTGGTGGAGTCCACAATTTCTTTAAAATGGAT
TTTGGCTGATTGCCCTAAAAGAGAGATCTGATAAAATGGCTCTTTAAATTTCTGTGAGTTG
GAATTGTCAGAATCATTTCACATTAGATTATCATAATTAAAATTTCTTTAGTTTCA
AAATTGTAATGGCTATAGAAAACAACATGAAATATTACAAATTGGCAACAATGC
CCTAAGAATTGTTAAATTCACTGGAGTTATTGTGAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCCATTTATTCTGGTCAATTATTGCTAGTGACACTGT
GCCTGCTCCAGTAGTCTCATTTCCCTATTGCTAATTGTTACTTTCTTGCTAATTGG
AAGATTAACTCATTAAATAAAATTGTCTAAGATTAAAAA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLDSTTLTSDESVKDHTAGR VVAGQIFLDSEEESEL
ESSIQEEEEDSLKSQEGESVTEDISFLES PN PENKDYE EEPKKVRKP ALTAIEGTAH GEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGF CETEEAAKRRQM QEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYL PONIQ AAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTGF GALGGN LIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCA GATTTAAGCCCATTCTGCAGTGGAAATTCTGAACTAGCAAGAGGACACCATCTTCTT
GTATTATAACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTCTGTGGACTCGTAAAGGAAA ACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGCTTGGAAACTTGGCAGCCAGAACCTTTGAT
AAAAAGGGATTCATGTAATCGCTGCCTGTACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTCTGGATGTGACCGACCCAGAGAACATGTCAGAGGACTG
CCCAGTGGGTGAAAGAACCAAGTTGGAGAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCGGCGTGTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCCCTGCAATCGTGGAGGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTCAAAACAAACTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGAGCAGCTGTCTCCAGACATCAAACAATATGGAGAACGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAGTCTCTCCCTAACAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTCTGGATAACCTCTGTCTCACATGCCAGCAGCTTGCAAGACCTTATTGGAAACAGAAA
GCAGAGCTGGCTAACCAAGGCAGTGTGACTCAGCTAACACAAATGTCTCTCCAGGCTATGA
AATTGGCCGATTCAAGAACACATCTCTTTCAACCCCATCCTTATCTGCTCCAACCTGGACT
CATTAGATCGTCTTATGGATTGCAAAGGGAGTCCCACCATCGCTGGTGGTATCCAGGGT
CCCTGCTCAAGTTCTTGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCCCT
GTATTTAGGCTTIGCCTGCTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCATTCAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACCTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEFYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAAGKDAKIFWIPLSHMPAALQDFLLLQKQELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTGGGCTTGTGCTCGGC
CACTCGCTTCCAGCACCTAACACGGACTCGGACACGGAAGGTTTCTTCTGGGGAAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATACAATTGA
CATTAGAAATATATTCCATGCTATCAGCTTTAGCTTATAATTCTCAGGGAAGTAAATG
AGCAAGCACTGAAGAAAATATTCAAAATGTCAAAAGAATGTTGGTAGGGTGGTACAAATTCCGT
CGTCATTAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAACTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAAACACCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAACCTCAAAAGGACTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGCATGCTGAACAAGGGTTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACAGCTCTAAATTTTGAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTTCTTGTCAAGCATTACGGACCTTTTCAAAATTCTGAATTCTCATTATGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACACTAACCCACATCTGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTGTTA
GATACACAAGACAAACGATCTAAAGCAAATCTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTGGTGAATATTACCGTCTC
CTACATTTGATCCTTTAACCTACAGGAGATTTTTATTGGCTGATGGGAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCAAGTTCTGTTACTATGTTCAAC
TGTTGCACTAATACACAGATAACTCTTAGTGCATTACTCACAAAGTACTTTCAACATCA
GATGCTTTATTCCAAACCTTTTCACTTCAAGTTGAGGGGAAGGCTACACAG
ACACATTCTTGAATTGAAAAGTGAAGCAGGACAGTGGCTCACACCTGTAATCCACGACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGTCATTAAAAAATGGAAAAGCAAGAAATGCTTCAAAATATGGAAA
GAAATTATATGAAAATTCTGAGTCATTAAATCTCCTTAAGTGAATCTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAATTTAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVGSCMSTGFSRAV
QTHSSKFFeedGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCALRTFFPNSEFLHSCVMSLKNRHSKSSCNYNHLDVVDNLTL
MVEHTDIPEASPASTPQI I KHKALDLDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGGAGGGCAGAGTCAGCCGAGCCGAGTCAGCCGGAGCAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCCAGCGAACGCCGCCGCCACACCCCTCTGCGGCTCCCGCGGCCCTGCCACCCCTCCCTCCCTCCCC
GCGTCCCCGCCCTGCCGGCCAGTCAGCTGCCGGTTCGCTGCCCGCGAAACCCGAGGTACCCAGCCGCCCT
GCTTCCCTGGGCCGCCGCCGCCCTCCACGCCCTCCCTCCACCCCTGGCCCGGCCCTGGCACCGGGACCCTGCCTGA
CGCGAGGCCAGCTACTTTGCCCGCGCTCTCCCGCTGCGCTCCACCAACTCCAACTCCTCTCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAA
GGTGGGAACGCGTCCGCCGCCGCCGACCATATGGCACGGTTGGCTTGCCGCCGCTTCTGCACCCGGCAGTGCTC
AGGCCGCCGCTGGCTGCCGAGCTCAAGTCAAAGTTGCTCGGAAGTGGCACGTCTTACGTGTCCAAGGCTTC
AACAGAACGATGCCCCCTCCACGAGATCAACGGTATCATTGAAGATCTGCCCCAGGGTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCTGCAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGCATCATTG
CAAGCTGCTTGTCTCACGTTACAAGAAGTTGATGAATTCTCAAAAGAACTACTTGAAGAATGCAGAGAACTCCCTG
AATGATATGTTGTGAAGACATATGCCATTATACATGCAAAATTCTGAGCTATTAAAGATCTTCGTAGAGTTG
AAACGTTACTACGTTGGAAATGTGAACCTGGAAGAAATGCTAAATGACTCTGGCTGCCCTCTGGAGCGGATG
TTCCGCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTGTGAGCAAGTACGGAGCAGCTGAAG
CCCTCGGAGATGTCCTCGCAAATTGAAGCTCAGGTTACTCGTCTTGTAGCAGGCCGTACTTCGCTCAAGGC
TTAGCGGTTGCCGGAGATGCGTGAAGCAAGGTCCTCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGCTCGTGAAGTGTGAACGGCATGTTACAACACTACTGCTCAAACATCATGAGAGGC
TGTGTTGCCAACCAAGGGATCTCGATTTGAATGGAACAATTCTAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTCACACATTGAATCGGTATGGATCCCCTCGATGTAAGATTTCTGATGCTATTATGAACATGCCAGGAT
AAATAGTGTCAAGTGTCTCAGAAGGTTTCAGGGATGTGGACCCCCCAAGCCCCCTCCAGCTGGACGAATTCTCGT
TCCATCTCTGAAAGTGCCTTCAGTGTCTCGCTTCAGACCACATACCCCCAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTTACTGATGTCAGGAGAAACTGAAACAGGCCAAGAAATTCTGGCTCCCTCCGAGCAAC
GTTTGCACAGATGAGGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGAATGGAAAGGCAAAGCAGGTAC
CTGTTGCAGTGCAGGAATGGATTAGCAACCCAGGGCAACAAACCCAGGGTCAAGCTGACACCAGCAACCCAGAC
ATACTGATCCCTCTGCAATCATGCTCTCGAGTGTGACCAAGATGAAAGATGCTACAAATGGAACGACGTG
GACTTCTTGATATCAGTGTGAAAGTAGTGGAGAAGGAAGTGGAGTGGCTGTGAGTATCAGCAGTGCCTTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCACATGAGAAAGCCGACAGTGTGCTGGTCCCTGGGCA
CAGGCCTACCTCTCACTGTCCTGCACTTGTGTTCTGGTTATGCAAGAGAGAGTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTCAATCAAAGTAAAGGCAACAGTTACCTTCTGACTGACTTGTGCTATTGCTTTTAAATGAA
TGGACAACATGTACAGTTTACTATGTGGCACTGGTTAAGAAGTGTGACTTGTGTTCTCATTGAGTTGG
AGGAAAAGGGACTGTGATTGAGTTGGTCTGCTCCCCAACCATGTTAAAGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTGTCATTGTGATTTTACTGCTGAGTTGGTCTGCTGGTCTTCTGCTGGCACGTAACAGTGTAGGTACAGAA
TCTGAAATATTAATAGCTGTACAGAAGCAGGTTTATTTATGTTATCTTATTAAAGAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELSKSKCSEVRRLYVSKGFNNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKS LNDMFVKTG
LYMQNSELFKDLFVELKRYYVGNNVLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLQLQVTRAFAARTFAQGLAVAGDVSKVSVVNPFTAQCTHALLKMIYCSHC
RGL VTVKPCYNCSNIMRGCLANQGDLDFEWNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIM
MQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLV
DK EKLKQAKKFWSLPSNVNCNDERMAAGNGNECDCWNGKGKSRYLFAVTGNGLANQGN
NPEVQV
DTS KPDILILRQIMALRVMTSKMKNAYNGNDVFFFDISDESSGEGSGSGCEYQQCPSEFDYN
ATDHAG KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypican proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCATTATATTCTTCAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGCGCCAAGAAATGTGAGTGAAAGATTGGCTCTGAG
AGCCCCGAGAAGAAAATTCAATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGTCTGCCCTTGTAGGAGCTCTG
AGCAGCCCACCTCTCCAATTAAACATTCTCAGCCAAGAAGACAGTGGAGCACACCTACCAGACACTC
TTCTTCTCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGTCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGCTCTCTAGTGTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCC
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPKCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

100 90 80 70 60 50 40 30 20 10 0

FIGURE 27

GGACGCCAGCGCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCACTCCTGCAGCTGCTGGTCTGCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGTCAGCCCTGTGCAAAGCTACTTCCCTACCTGA
TGGCGTGTGACTCCAAAGAGCAACCGAAGATGGAGAGCAAGAAACGGAGCTTCAAGCCAG
ATAAAGGGCTTACAGGAGCTCCGGAAAGTGGCCTACTGGAGCTGGCTGCGAACCGGAGC
CAACTTCAGTTCTACCCACCGGGCTGCAGGGCACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGTG
CTCTGTGCAGGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTCTGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCACCTGAAACACATTGGGATGGCTGCTGCCCTACCAAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCCGAAATCAAATGGAACGACAGCCCCCTCCCTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGAAAGGCTGCAAAACATTCTCCCAAGCTCCAAGGCACACTTGTCC
TTCTCCAGCCTCAATTAGAACAAAGCCACCCACCAGCCTATCTATCTCCACTGAGAGGGACTA
GCAGAATGAGAGAACATTCATGTACCCACCTACTAGTCCCTCTCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCGCCTCGACAGTGAAAAAGCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCTCAACTGCAAGTTCTGGACTAGTCTCCCAACGTTGCCTC
CCAATGTTGTCCTTCCCTCGTCCATGGTAAAGCTCCTCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACGTGGTCAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCACTACCACCTTCTCCTGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTAATAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSI
QLEQATHQPIYLPRLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCTTACCTATGCTGCTGCTAACGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTTGGAGTGGGACTGGTCGGTGCCAGAAAGTCTCTCTG
CCACTGACGCCCCATCAGGGATTGGCCTCTTCCCCCTCCCTGTGTCTGCCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGAGATAGGGAACAGAAGAGGGTAGTGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTTTATGATTCTTACTAATTATACAAAGATATTAAGGCCCTGTCATTAAGAAATT
GTTCCCTCCCTGTGTTCAATGTTGTAAAGATTGTTCTGTGTAAATATGCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTCCTGAACTTCCCTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCCTTATTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTGATAAATGATGTTGACACCCCTCACCGAATTCTAAGTGAATATGTCGG
GAAGAGATAACAATCCTGGCCTGTATCCTCGATTAGCCTTGTCTTGGCCATGATGTTTACC
TTCAGATTTCATCACCAACCCCTCTGGTCACATTTCATTGGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGATTATGACTATAACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTGGCTGCTGGGTTGCTATCGTATCCACAGGCATCAGGCAGTG
CTGCTCGTCTGATTTGTTCTAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGGCTCTGGGCGCTGTGCTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCAGGCAAGTGAATATAAGCCCCTTCGGCATTGGTACATGTGGTGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCAGCAAATGACTATAGCTGGGAG
TGGTTACTTGTATTCAACAGAAGTAAAATGATCCTCCTGATCATCCCATCCTTGTCTCTC
TCCATTCTCTCTTACCATCAAGGAACCGTTGTGAAAGGGTCAATTAACTCTGTGGTGTGAG
GATTCCGAGAATATTGTATGTACATGCAAGCAGACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTCGATGCTGACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACAGAATGCATATACTACAACTGCTATTAACTGGACAGATTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTGTCCAAGAACTCAAGTCACCTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATTGGTAGCTTTTGCCACTTAGTAC
CCATAGTTTTATCTGTGTTGAAACTGTCAGTGTGCTGGATGCACCTTCTGTGTTGCTGTC
TGGAAACAAATGATGGATCGTCAGAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCAGTCAATTAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGAATGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGACCCTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATTAAACCCATTCTCTCTCAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVLGGLFVCGLWWLYDYTNLDSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMVMQNALKEQQHG
ALSRYLFRCYCFCWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
IIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTCCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAAGCCTAAGGGAAAAGAATATTCTGTGTTGGT
GAAAATTTTGAACCAAGGGTGTCAATTCTGCTTCAAACAAAGGGTGTCAATTCTGATATTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTTCTTGTGTTGACTGGAGTAC
ATTCAAACAAAGAAACGGCAAAGAAGATTAAAGGCCAAGTTCAGTGTGCCAGATCACTGC
GATGTCAAAGCCGGAAAGATCATGATCCTGAGTTCAATTGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGCCACTGACGGTGTATGCATCCTACTCCAGTGTGTTGGCGCTGCCG
TACACAGTGGTGTGCTTGATAATTCAAGGAGGAAAATACTTGTGTTGGAGGTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCGATGGAGAGAAATCCTT
TATCGTCTTAGAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTTACATACTCATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCAAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACGTGACACGGCGGTCACTCTGATGCAGCTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCCTCTGCTGCTTCTACCCACAGCATCCCCAGACCCACAATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAACAGGCC
AGAGCTGATCCAGGTATCCAAGGCAGATCTCAGGAGCTGCCCTCCAGAAACCTGTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGATTGAGCACACAGTCTTGGAGCCAGTATCCC
TGGGAGATCCAAGTCAAAATTGACTTGTGTTTAATTGATGGGAGCACCAGCTTGGCAA
CGGCAGATTCCGAATCCAGAACAGCAGCTCTGGCTGATGGTGGCCAGCTTGACATTGGCCCTGC
CGGTCACACTGATGGGTGTTGTCAGTATGGAGACAAACCTGCTACTCACTTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTCTAAT
GTAGGTGGGCCATCTCTTGTGACCAAGAACCTTCTTCCAAAGCCAATGGAAACAGAACGG
GGCTCCAAATGGTGGTGGTGGTGGATGGTGGCTGGCCACGGACAAAGTGGAGGAGGCTTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTCATCACCATTGAAGGTGCTGCTAAAATGAG
AAGCAGTATGTGGTGGAGCCAACTTGCACAAACAGGCCGTGTCAGAACAAACGGCTTACTC
GCTCACGTGCAAGAGCTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCCTGGCCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCGTATCGAC
GGCTCCAGCAGTGTGGGAGGGCAACTCCGCACCGTCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGGCATCGGGCCGTGCACTACACCTACGAACAGCGGC
TGGAGTTGGGTGCAAGTACAGCAGCAAGCCTGACATCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTTGGCACCAGCACGGGGCTGCCATCAACTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAAGAGGAAGTTATGATCCTCATCCGCACGGGAGGTCTACGACAGCTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCCATAGGCCTGGCTGGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCAACTTGTACAGAGTTCAACTCAC
GTTTGACAACCTCCATCAGTATGTCCCCAGGATCATCCAGAACATTGTGACAGTGGAGGAGT
AGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAAGTGTGCTTTACTAAGTACGTGTT
GGACCAACCCACCGCTTAATGGGAGCCACGGCATCAAGTCTGGGAGGCTGGAGAAAC
AAATGTCTGTTATTATTCTTGCATCATGCTTTCTATCCAAAATGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAGGAAAGGCTACATCATGTTGAGGGTGTGGAGAATTTCAT
TTTGACAATTGTTCTAAAATGGTGGAAATACAGTGCAGCCCTACGACAGGCTACGTAG
AGCTTTGTGAGATTGTTAGGAAATTGCTGAAATTGTTAGAAGGATGAAAATAAAAAA
AAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLSEPKKKGVTPSALTYSSSKPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPPTTLPDSPSAASSTTSIPRPPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSSLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPILMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMDGWPDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQPLVKRVCNDRLACSKTCLNSADIGFVIDGSSSGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGYWSGGTSTGAIAINFALQL
FKKSCKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTCTGTTTCAGAGTAGTCACAACAGATCTGAGTGTAAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCTATCTGAAATTCCACAGTTCTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTCTATCACGTGGTGCCTCCGACTACTCACCCGAGTGT
AAGAACCTTCGGCTCGCGTGCTCTGAGCTGCTGTGGAGGCCTCGGCCTCTGGACTGTCCCTCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCGCTGCTGTCACCTCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTGAACGGATGTACTTCTATGAGTATGAGCCGATT
CAGACAAGACTTCACTTCACACTTCGAGAGCATTCAAACAGCTCTCATCAAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCCTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGT
TGGTGGGATATGAGGTTCTACATTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGG
ATTGTCCTAGAGGATGAACACCTCTTATGGTGACATAATCCGACAAGATTAGACACATATAATA
ACCTGACCTGAAAACCATTATGGCATTCAAGGTGGTAACTGAGTTTGCCTTAACTGCAAGTACGTAATG
AAGACAGACACTGATGTTTCAATACTGGCAATTAGTGAATCTTAAACCTAACCAACTCAGA
GAAGTTTCACAGGTTATCCTCTAATTGATAATTATTCTTATAGAGGATTACCAAAAAACCATATT
CTTACCAAGGAGTATCCTTCAAGGTGTTCCCTCCATCTGCAGTGGGTGGTTATATAATGTCCAGAGAT
TTGGTGCAAGGATCTATGAAATGATGGTCACGTAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAGGATGAACATTCAATTCCAGAACAGACACAAATCTTCTTATATAGAATCC
ATTGGATGTCGCAACTGAGACGTGTGATTGCAGCCATGGCTTCTTCAAGGAGATCATCACTTT
TGGCAGGTGATGCTAAGGAACACCACATGCCATTTAACTCACATTCAAAAGCCTAGAAGGACAG
GATACCTGTGGAAAGTGTAAATAAGTAGGTACTGTGAAAATTCAATGGGAGGTGAGTGTGCTGGCTT
ACACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTCAAGATGATATGTGGAGGAATTAAATAAGGAATTGGAGGTTTGCTAAAGAAATTAAAGG
ACCAAAACAATTGGACATGTCATTCTGTAGACTAGAAATTCTTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAAACAATGTAGAGTTTATTGAAACAATGTGACTTGAAGGTTGTGTA
TATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTGTCAAAAACTCTTCACTGAAGTT
CTGAACAAAATTTCACCTGTTTGGCATTTATAAGTACTTCAGACTAGATGTCAGTATTTCACAGTTATT
ATTATTTAAAATTACTTCACCTTGTGTTTAAATGTTTGACGATTCAACACAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACGTGATCAGTTATTATGATAACATCACTCCA
TTAATGTAAAGTCAGGTCAATTGCATATCAGTAATCTTGGACTTGTAAATATTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNC SHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDII RQDFLDTYNNLTLKTI MAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCGLGYIMSRDLVPRYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLLISLSPREYTVINEACPGAEWNIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRCVGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPIKACREPDKISDLVRRRLPQMVSRETPLHOLYSAAFSKQKLQSAPTK
KPALPFGDLPMPGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENTAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDRDEKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFAQESHITVAGWNVLADVRSPGFKNDSLRSGVSVVDSLLCEEQHEDHGI PVSVDNMFCA
SWEPTAPS DICTAETGGIAAVSFPGRASPEPRWHLMLVWSWSYDKTCSHRLSTAFTKVLFFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTCCTACATCCTCATCTGAGAACAGAGACATAATCTTACGGGCCGTGATTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTTAAAGGAGCTGGCTGG
TTGGGCCCTGTAGCTGACAGAACGGTGGCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTGTG
TGGCTTGCCCTGGCTCAGTCTGCTAACTACATTGACAATGTGGCAACCTGCACCTCTGTATTCAAACACTGTA
AAGGTGCCTCCACTACGGCTGACCAAAGATAAGGAAGAGGCCTCACAAAGATGGCTGCCAGGGCTGTGCCAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTCTGCAGCTGCCACCATCTCTTAATGACAGACGGCTGGCTAGACA
ACCTGCCTACGTGTCTCGCAGAGGACGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAAGTA
GGCACGGCCCTTGAGAGATCCACTATTAGAACAGATCAATTAAAAAATAATCGAGCTTGAGTGTCTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCAACCATGCCGACCAGGGCAGGGAAATTCTGAAAACACACTGCCCTGAAG
TCTTCCAAGGTTGTACACCTGATTCCAGATGGTGAATTACACCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCCTCTATTAGGTGGTGGAGGTAGCGAAACCCACTGGTCCATATCATTATCCAACACATTATCGTATGGG
TGATGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCACCGGATGGACATCAGCAATGCCCTC
ACAACATACGCTGTGCGTCTCTGGCAGCCCTGCCAGGTGCTGTGCTGACTGTGATGCCGAAAGAAGTCCGCA
GCAGGAACATGGACAGGCCCGGATGCCCTACAGACCCGAGATGACAGCTTCATGTGATTCTAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAAACTGGTCCGCAAGGTGGATGCCCTGGGGTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAAGTGGCGCTCATCTGATTCAAGGCCAGTGAAGACGCTGTCACCTCGTGTGCCCCAGGGTCCGGCAGC
GGAGCCTGACATCTTCAGGAAGCCGGTGGCAACGCAATGCCAGCTGGTCCCCAGGGCAGGGAGAGGAGCAACAA
CTCCCAAGCCCTCCATCCTACAAATTACTGTGATGAGAACGGTAAATATCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTGCAGGGGAGCATCACATAGAGAACGGGATTGCTTATCTATGTCATCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAAGAACAAAAACAGGTGACATTGTTGAATGTGGATGGGTGAAACTGACAGAGGTGAGC
GGAGTGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTGGAAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCCACGCCCCACTCCAAACACATGGCCCCACCCAGTGAAGTGGTCCCCATCCT
GGGTCATGTGGCTGGAATTACACCGTGCTGTATAACTGTAAGATATTGTTACGAAGAACACAGCTGGAAGTC
TGGGCTCTGCATTGTAGGAGGTTATGAAGAACATGGAAACAAACCTTTTCATCAAATCCATTGTTGAAGGAA
CACCAAGCATACAATGATGGAAGAACATTAGATGTGGTATATTCTCTTGCTGTCAATGGTAAAGTACATCAGGAATGA
TACATGCTTGCTGGCAAGACTGCTGAAAGAACATTAAAGGAAGAACATTACTCTAACTATTGTTCTTGGCTGGCACTT
TTTATAGAATCAATGATGGGTAGAGAACAGAACAGAACATCAAATAGGCTAAGAAGTGAACACTATAATTATC
TTGTCAGTTTATATTAAAGAACATGTCAGGGAAACAGAACAGAACATCAAATGTCAGGAAAGTATGTCATCTAAAGGCCAGTT
ACACCTCAGAAATATGATTCCAAAAAAATTAAACACTAGTTTTTCAGTGTGGAGGATTCTCATTACTCTAC
AACATTGTTATATTCTATTCAATAAAAGCCCTAAACAACTAAATGATTGATTTGATACCCACTGAATT
CAAGCTGATTAAATTAAATTGGTATATGCTGAAAGTCTGCCAAGGGTACATTGGCATTAAATTACAGCT
AAAATTTTAAATGCATTGCTGAGAACGTTGCTTCATCAAACAAAGAATAATTTTCAGAAGTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNQAPDAYRPRDDSFHVLNKSSPEEQLGIKLVRKVDEPGVIFFNVLDDGVAYRHG
QLEENDRVLAINGHDLRYGSPEAAHLIQASERRVHLVVSQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTGCAATCAGATTGGAAAAGCTCAACTTGAAGCTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACCGTAATAAAAACATGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGCTGACAGTGGTTGGGTGGC
CACCACTAACTACTTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAGGGAAAACCTCTGACTAATGAAGCATTCCACGAAGAAGGTAGAACTT
GACAACGTGCTCTGTCTCCTTACCTCAGAGGCCAGCAAGCTCATTTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGAACAGAGAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCAAACCTTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACGTACAGTGGATATTGG
GGGTGTTACTGCCCTAACGAGAGAGCAGTTTCAAGGTGAATGGATTCTAACAACTACTGGGAT
GGGGAGGCGAACAGCATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCGGCCCTG
CCTGAAGTGGTAAATATAACATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
GATGAAGCTCTAACCAAGTGTACGAGTCTGGAGAACAGATGGTGAAGTAGTTGTTCTTAAAT
TAGTATCTGTGAAACACAATCCTTATATCAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGAATGGAAAGAACTGATTCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGGTGAAGTGAATTCTTGTGTTCTTGTATT
TAGCAGAGCTCTGGTGAAGTAGTATAAAACAGTTGTAACAAGAACAGCTTCTAGTCATTGAT
CATGAGGGTTAAATATTGTAATATGGATACTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTTAAATTGAAAGTAAATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGCTGAGAGAACCCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCACTTACAGTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
GAGAAGGCGTCCACAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGAGGACCTGCCAGGTATGCCCTCAGTGTGATGCCACAGAGAACATCTCTATTAGT
TTTAAAGAGTTTGTAAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAACCTCTGTAGTAAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLGKGKTLN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNNGFSNNYWGCGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTGCCAGCCTGACCAGTGGCTCTGTTTCCCACAAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGGCCAGGGCCAGCTGGATGCCAT
GTTCCAGAGGCAGAGGAGGCAGACACCCACTTCCCCATCTGCATTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCCCTCCC
TTCCTTATTTATTCCCTGCTGCCAGAACATAGGTCTTGGAAATAAAATGGCTGGTTCTTGTT
TCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLILASLTGSVFPQQTGQLAEIQLQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGGTGGTGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTCAACACAACCCCTTTGTACCCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAAG
AAGAATGACTCAGGGATCTACTATGTGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCTGTCAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCCTCCTGTGTCCTGTTGGTGCCCTCCTGCTCAGTCTTTGACTGGGCTATTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTGTCGGG
AAACTCTAACATATGCCCTCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYIILWQLTGSAAAGPVKELGVSGGAVTFPLKSVKQVDSIVWTFNTTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPIIARKLCEGAADDPDSSMVLLCLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATAACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAACCCCATCTTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGCTGATGGCATTCCAGCAACAACATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAAGGTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTAATTTCATGAAAAACATCAGTGACATTCCAGAATCCTCAA
CTTGCAGTGGTTTCCAATGACTCTTGCACCTCTACTGGTTCAATAAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTTTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGCT
CAGTCAGATAGTCATCGGTTCTGGCTGTGTGGAGTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGAATAAAATGTAAGTATCAGTAGTTGAAAAA

FIGURE 48

MTCCGTSCNGFSLVLLLGVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTTCTCTACTGGCAGGCCTGACTGCCTTGAAGCCAATGACC
CATTGCCAATAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCTGCCATTGCTGGGATCGCGGAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACAGCACAGTCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGCCCTGAAGCCTAACACTGGCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGCACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGCAAGG
AGGAGACCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCTGGCTGCCCTGCTGGCCCTCTGCCCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGACAAATATTGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGGTCGGGAAGCAGCCATGCTGGAAACACTGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGACAGATGCTGTCCGGCTCTGGCAGGGG
GTGCCTGGCCACAGTGGTCTGGAAACTCTGGAGGCCATGGCATTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAAGGCCATCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGAAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGCAGCAG
AATGGCAGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGGAGGTACAGCAGCAGTGGAGT
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGGAGGTACAGCAGCAGTGGAGT
CCTCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAGCCAGGGAAATGAAGCCCGGGAGCGGGGAATCTGGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGGGCAATGCCCTTGG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCAGCTGGGAGTGGAGGAGGTACGCTGTT
GGTGGAGTCATACTGTGAACCTGAGACGTCTCTGGATGTTAACCTTGACACTTTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGATGCCATAAAACAGGACCAAGAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGCCCCCACACTCCCTCTAA
AACACCACCCCTCATCACTAAATCTCAGGCCCTGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AA

FIGURE 52

MKFQGPLACLLLALCILGSGEAGPLQSGEESTGTNIQEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGQSQGGLGGQQGNPGGLGTPWVHGYPGNSAGSGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPSGSGGSSNSGGSGSQSGSSGSGSNGDNNNGSSGGS
SSGSSSGSSSGSSGGSSGSSGSSGNSGGSRGDGSESSWGSSTGSSSGNHGGGGNGHKPGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGGSGDNYRGQSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFNWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGGACAAGCTGCTCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGCCCCATGGCTACTCCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACTGGTTGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCC
CATCCCCTCATGTTTATGCCACCCCTGACACCATCCGGTCTATCACCAATGCCTCAGTGCCA
TTGCACCCAAGGATAATCTTCATCAGGTTCTGAAGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCGCCCTCCATTCAACATC
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCAC
CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGCACATCAGCCTCATGACCTGGAGCAGTCTA
CAGAAATGCATCTCAGCTTGACAGCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCTTGTAGAGAAAAGCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGGGCGCTTCCACAGGGCTGCCCTGGTGCATGACTCACAGAC
GCTGTCATCCGGGAGCGCGCTGCACCCCTCCCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATGATGTCGCTTGAGCAAGGATGAAGATGGAGG
CATTGTCAGATGAGGATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCCTGGTCTGTACAACCTGCGAGGCACCCAGAATACCAGGAGCGCTGCC
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCATTCTAAAGAGATTGAATGGGACGACCTGG
AGCTGCCCTCCTGACCATGTCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCAGTGCACCCAGGACATTGTTCTCCAGATGGCGAGTCATCCCCAAGGCATTACCTG
CCTCATGATATTATAGGGTCCATCACACCCAACTGTGTGGCCGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTATTCCCTTCAGC
GGGCCAGGAACGTGCATGGGCAGGCAGTCGCACACTGAGCCCCCGCAGGAAGCTGGAATTG
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGCAGACTTCTGAC
CCATCCACCTGTTTTGCAATTGTCATGAATAAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWGLRPVAMSPWLLLLVVGSWLLARILAWTYAFYNNCRRLQCFCQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEIGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRFHRAC
RLVHDFTDAVIREERRTLPTQGIDDDFKDKAKSKTLDFIDVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPVEYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCCTCATGGGACCACTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACCTACCCGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACCTATCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGGCCAGTTTATGAAGCTTGAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGTCTTACAGACATGTGCCTTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCCCTC
TCTGGATGTTGTCCTGAATTCCCATGAATAACAAACCTATTCAACAGCAACAGCAAAAAAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCGAATCGGCTCGAGGGGAGTGGAGCACCCAGCAGGCCAACATGCTCTGTGCGCTG
TACCGTGCCTCATCGGGAGGCCAGACCGAGTCCAGTACTTTGACTCGAAGGGCTCCCTGCCAGCTGAAGTCC
ATTTCAAGCTCAGTGTCTCATCCCCCTCCAGGAATTCTCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTGAAGAATTGTCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGATTGGACAAAAGAATGATGGACGATGAGCCAGGGAGATCATGAGTCCCTGCCAGCTTG
GGAGTCAGAGATATCTGAACAGCAGGCAGAAAAATTCTCAAGAGCATGGATAAAACGGCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCAACCTCTCCACCCCGTGGAAAACATCCCCAGAGATCATCCTCTACTGAGAGCATCCACG
ATCTTGATGTGGTGGAGAACTAACGGTGGGATGAGTCACTGGAGAGGAGACAGGGGGATGTGGTGGAGA
CACCTGGTGGCAGGGAGTGGGGCAGGGGGCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCATGCCCTCCCGCAGAACACATGGCATGTTGGCTTCACTCAGATGATTGAGAAGGAGGGCCAGG
TCACCTGGCGGGGAATGGCATCACGTCCTCAAATTGCCGAATCAGGCAATCAAATTGCGCTATGAGCAG
ATCAAGCGCTTGTGGTAGTGCACAGGAGACTCTGAGGATTACAGAGGGCTGTGGCAGGGCTTGGCAGGGCC
ATCGCCAGAGCAGCATCACCAATGGGCTCTGAAGACCCGGATGGCCTGCCAGAACAGGGCAGTACTCAGGA
ATGCTGGACTGCGCAGGAGGATCTGGCAGAGGGGGCTGGCCCTCTACAAGGCTATGTCCCCAACATGCTG
GGCATCATCCCTATGCCGCATGACCTTGAGTCACTCAGAGACGCTCAAGAATGCTGGCTGCAGCACTATGAGCTG
AACAGCGCGACCCCGCTGTTGTGCTCTGCCGTGGCACCAGTCCAGTACCTGTGGCAGCTGGCAGCTAC
CCCCGGCCCTAGTCAGGACCCGGATGAGGCGCAAGCCTCATGGGCTCGGGAGGTGACCATGAGCAGCCTC
TTCAAAACATATCTGCCGAGGGGGCTTGGGCTGTACAGGGGCTGGCCCAACTTCAAGGGTCAATGAGGTCATCCCA
GCTGTGAGCATCAGTACGTGGTACAGAGAACCTGGCGTGGCAGGCTGAGCTGGCGTGTGAGGGGGAGGGC
CGCCGGACTGGACTGCTGATCTGGCCGAGGCTGGGGTGTGAGCCATCTCATTCTGTGAATGTGCAACACT
AAGCTGTCTGAGCCAAGCTGTGAAAACCTAGACGACCCGAGGGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCTGCTGACCCAGACCCCTCTGTGGTCAAGCGAAGACCCAGGCAATTCTTAGGGTCCAGGGTCAAGCAGG
CTCGGGCTCACATGTGTAAGGACAGGACATTCTGCTGAGTGCCTGCCAATAGTGAGCTGGAGCCTGGAGCCGGCT
TAGTTCTCCATTTCACCCCTTGCGACGGCTGGCCACGGCCCTGGCTCTGGTCTGCCGTGCACTCCTGTGC
CCTCTGCTGCTGCCGTGAGGTAAGGGGGAGGGCTACAGCCACATCCACCCCTGTCATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCTCCACGGCTGACTTCCAACTACAGCATGACGCCAATTGGC
TGTGAGGAAGAGGAAGGGATCTGCCATTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGCACTG
CTGGGAGTGCAGGGGGCTGGGCTGCCATTGGCAGCAGGAAGGCAACTGGGCTCATGGTCTGAGCT
GGCCTGGACCCACTGTCAGGATGGGGCCACCTCAGAACCAAACACTGTCCTCACTGTGCACTGAGGAGCA
CCATGTTGAGGGCGAAGGGCAGAGCTGGTGTGTTCTGGGGAGGGAAAGGAAAGGTGTTGGAGGCTTAATTATGG
ACTGTGGAAAAGGGTTGTCCAGAAGGACAAGCCGACAAATGAGCGACTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTGGTCACTGCTCAGGTTGCTGACGCCCTGGGGTCTGCCACCCAGCAGGGGGCAGC
GGGACAGCCACATTCACATTGCTGCACTGGCAGGACCTATTGTTATTTGTTATTTGAACAGAGTTATGTCC
AACTTTTATAGATTGTTAAATAGCTGTCATTTCAGGTTCACTTATTTATCATATTATGTTATGGT
GATTGACCTTCCAAAGCCGCCAGTGGGATGGAGGAGGAGGAGAAGGGGGCTGGCCGCTGAGTCACATCT
GTCCAGAGAAATTCTTTGGGACTGGAGGAGAAAAGCGGCCAGAAGGAGCCAGGCCCTGGCTCCCTTGGCAG
GTTGGGAAGGGCTGCCCTAGGATTCAAGGGTTGACTGGGGCTGAGAGAGAGGGAGGAACCTCAAT
AACCTGAAAGGTGGAATCCAGTTATTCTGCGCTGCCAGGGTTCTTATTTCACTCTTCTGAATGTCAAGGCAG
TGAGGTGCTCTCAGTGAATTGTGGTGGGGGGCTGGAGGGAGGGGGCTGGCTCCCTCCCAGC
CTTCGCTGCCCTGCTTAACATGCCGCCAACCTGGCAGCTCACGGTTGCACTTCAATTCCACAGAATGACCTGA
TGAGGAATCTCAAGGATGCAAGATCAATGCAAAAATTGTTATATGAACATATAACTGGAGTCGTC
CAAATAAGAAGAATTGGACGTTAGAAGTTGTCATTAAAGCAGCCTCTAATAAGTTGTTCAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFOYFESKGLPAELKSIFKLSVIFPSQEFSTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMKDNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKSTIFDVGENLTVDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGETQMIREGGARSILWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNawlQHYAVNSADPGVFVLLACGTMSSC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACTGGCTGGGACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGAGGATGGAATCCTGAGCTGCACTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTTGGTCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTCAGAGCCGGACAGCAGTGTTGCTGATCAAAGTGTAGTT
GGCAATGCCTTTGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGCTGGGCATCCCAAGTTGACCAAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGCTGCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGT
ACAGAATCGGAGATCAAAGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTT
CTCTCTTCTTGCCATCAGCTGGGCACCTCTGCCTCTCAGCCCTTACCTGATGCTAAATAAT
GTGCCTGGCCACAAAAAGCATGCAAAGTCATTGTTACAACAGGGGATCTACAGAACTATTCAC
CACCAGATATGACCTAGTTATATTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGCAAGGCCAAAGCAGAAGGCTCAATATGAACAAGATAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTGATGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTCATCCCCAGATCTCAGGGACCTCCCCGCTGT
CACCTGGGAGTGAGAGGACAGGATAGTCATGTTCTGCTGAAATTAGTTATATGTC
TGTAATGTTGCTGAGGAAGCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCCTAACAGCCTGCTAATTGACTGCCACTCCGCAACTCAGGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCCAACTGACAATGCCAAAGTTGAGAAAATGATCATAATTAGTCATAACAGAGCAGT
CGGGGACACCGATTTATAAAACTGAGCACCTCTTTAAACAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVWWASQVDQGANFSEVS
NTSFELNSEVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKVTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMNK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAACATCACCTGGCCAGCTATCCTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCTGGTGGTTATGGGGTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGATGTTCCCTCTGGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAGCCACCTCCAAGTTCTACGGTGCCAGCAGCCT
GGGCTTATGGACAGGGTGGCGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCATGATGATAAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTCTCAGCCCTGTGGAAATTCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTCACCCAGCTCTGGCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTGACCGCTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTTGACCAACCATCTGTGGAGAGTGGAGTGCACCAGGACCTT
TCCTGGCTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AATAGTGAGGACCGGGCTGAGGCCACACAGATAGGGGCTGATGGAGGAGAGGATAGAAGTGA
ATGTCCTGATGGCATGAGCAGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTGAATGTTCATGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCATCTGTCAAACGAGCCATTCTCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGCCCTGCTAGACAAATCTGCTCCCTGGCATT
CTTGGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTGCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCAATTTC
ATTGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYQQGGAPPNDPEAYSWFQSVDSDH
SGYISMKELQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSAWKFQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLOVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGG**A**TGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTGC~~CCGGAGCATCGTGTGAGGAGAAAGTTCC~~AAACTTCGGGACC
AACTTGCCTCAGCTCGGACAACCTCCACTGGCCCTCTAACTCTGAACATCCGCAGCCC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCCCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGCTGCCTGCCATG
GATTCCCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCCTGGGG
AGCGCTGCCTGAAGAACTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGCAGTGGCCCTT
TGCCCTGGGAGTCTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAATTCACTGGAGCCGGGGAAAAATCCTTCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTCTGCCTGATCACCCCTGGGTACCCGAATCCAGTG
TGTCCCTGGGAGGTGGAGGCCCTGGACTGGTGGGAACGAGGCCATGCCACACCTGAGGGA
ATCTGGGTATCAATAATCAACCCCAGGTACCAGCTGGGAAATTAAATCGGTATCCAGGAGG
CAGCTGGGAAATTAAATCGGTATCCAGGAGGCAGCTGGGAAATTAAATCGGTATCCAGGAG
GCAGCTGGGAAATTACATCTATACCCAGGTATCAATAACCCATTCCCTCCTGGAGTTCTCCGC
CCTCCTGGTCTCTTGGAACATCCCAGCTGGCTCCCTAATCCTCCAAGCCCTAGGGTGCAGTG
GGGCT**A**GAGCACGATAGAGGGAAACCAACATTGGAGTTAGAGTCCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCACCTTTCAGTGCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEEELSYLSSAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMHPPEGIWGINNQPPGTSWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPVGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGCGCAGTGGCCACTATGGGT
TGGGCTGCCCTGTCCCTCTTGACCCCTGGCAGCTCACATGGAACAGGGCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTGACAAATTCTCTATGAGTCCAGCTTCTGGAA
TTGCTTAAAAGCTGCCTCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTAACACCATGTTGTCGAAACACATGACAGCCATTGAAGCCTGTGTCCTTCTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGGCCCCGACCCGTCTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLGGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATGCCGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGGCCTGCCAACCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGGCCAGCCTGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCAGTGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTCTGGCTGCCTTGCTGTGGCTGTGGGCTG
GCACGTGTTGGTGCAGCTCAGGGAGAACGGCAGACGGAGGGCACCTACCCGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCATCTAGGTCCCCCTCTGCATCTGTCTCCCTATTGCTGTGACCTTGGGAAA
GGCAGTGCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLLLLALGLPFLARWGRAGQIQTTSANENSTVLPSSSSDGNLRPEAITAIIVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAAATACTAGAGAGGAACAAATGGGTTATTCAAGGGTTTCTCTAGTCTGCTGCCAG
TCAAATACTCCTCATTAAGCTGAATAATAATGGCTTGAAGATATTGTCTAGTCTAGTGTGCCAGAA
GATGAAAAAAATAATTGAACAAATAGAGGATATGGTCACTACAGCTTCTACGCTTCACTGTTGAAGGCCACAGAAA
TTTTTTTCAAAATGTATCTATTAACTCTGAGAATTGGAAGAAAATCTCAGTACAAAAGGCCAAACATGAA
AACCATAAACATGCTGATGTATTAGTCGACCCACTCACTCCAGGTAGAGTGAACCATACACCAAGCTTCA
GAATGTGGAGAGAAAGGCCAAATCATTCACTTCACCCCTGACCTTCACTTGAAA
ACAAAATGAATATGGACCA
CCAGGCAAACTGTTGTCATGAGTGGCTCACCTCCGGGGAGTGTGTTGATGAGTACAATGAAGATCAGCCTTC
TACCGTCTAAGTCAAAAAAATCGAACAAAGGTGTTCCGCAAGGTATCTCTGTTGAGAATAGGTTTATAAGTGT
CAAGGAGGGCAGCTGCTTAGAGCAGTCAGGAAATTGATTCTACAAACAAAATGTTGAAAGTGTGCAATTCTT
CCTGTAAAACATACAAACAGAAAACATCCTAATGTTATGCAAGGATTATTGATTCTGTTGATTGTAACGAA
AAAAACCCATAATCAAGAAGCTCAGCCTACAAAACATAAGTCGCAATTGAGAATCAGGAGGTGATTGCAAT
TCTGAGGATTAAAAACACCATACCCATGGTGACACCACTCCTCCACCTGTCTCTCATGCTGAAGATCAGCAA
AGAATTGTGCTTAGTTCTGATAAGTCTGGAAGCATGGGGGGTAAGGACGCCCTAAATCGAATGAATCAAGCAGCA
AAACATTCTCTGTCGCAACTGTTGAAATGGATCTGGTGGGGATGGTCACTTGATAGTACTGCCACTATGT
AATAAGCTAACCTAAACAGTGTGAGAAGAACACACTCTGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTGGATTAAATATGCACTTCACTGGGAGGCTACATTCTGGGAGGCTACATTCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGGAGGATAACACTCAGGTTCTGTATTGATGAGTGAAGAACAAAGTGGGCCATTGTCAT
TTTATTGCTTGGGAAGAGCTGCTGATGAGCAGTAATAGAGATGAGCAAGATAACAGGGAGAAGTCATTGTTATGTT
TCAGATGAAGCTCAGAACATGGCTCATTGATGCTTGGGCTCTTACATCAGGAAATACTGATCTCCAGAAG
TCCCTCAGCTCGAAAGTAAGGATTAAACACTGAATGCTGATGAGCAGACTGTCATATTGATAGTACA
GTGGGAAAGGACACGTTCTTCATCAGAACACTGCTGCCCTCAGATTCTCTGGATCAGGAACTGGCAAC
ATAATGGAAAATTCACAGTGGATGCAACTTCCAAAATGGCCTATCTCAGTATTCTCAGGA
ACTGCAAGGTGGGACT
TGGGCATACATCTCAAGGCAAGCGAACCCAGAAACATTAACTATTACAGTAACCTCTCGAGCAGCAAATTCTTCT
GTGCTCCAATCACAGTGAATGCTAAAATGAATAAGGACGTAACAGTTCTCCAGGCCAATGATTGTTACCCAGAA
ATTCTACAGGATTGTACTGTTCTGGGAACTTGTGAGCTGTTCTCATTGAATCAGGAACTGACAG
TTGGAACTTTGGGATAATGGTCAGGGCTGATTCTTCAAGAATGATGGGACTCTACCCAGGTT
ACAGGAAATGGCAGATAATGCTTAAAGGTTGGGCTCATGGAGGACCAAAACACTGCCAGGCTAAATTACGGCTTCCA
CTGAATAGAGCCCGTACATACCAGGGCTGGTAGTGAACGGGGAAATTGAAGCAACCCGCCAAGACCTGAAATTGAT
GAGGAACTCAGACCACTTGGAGGATTCTAGCCGAACAGCATTGGAGGTGCTTGTGATACAGT
CTTCCCTCTGCTGCCAACATACCCACCAAGTCAAAATCAGACCTGGTACAGGATCACTGAGGAAAGTATTCT
ACATGGACAGCCAGGAGATAATTGATGGTGGAAAAGTCTCAAGCTTATATCATAAGAATAAGTCAAGGAT
GATCTAAGAGCAGCTTGTGATGCTCTCAAGTAATTAACACTGATCTGTCACCAAAGGAGGCCACTCAGGAA
AGCTTGTGATTTAAACAGAAAATATCTCAGAAGAAAATGCAACCCACATATTGCTTGCATTAAAAGTATAGATAAA
AGCAATTGACATCAAAGTATCAACATTGACAAGTAACCTTGTGTTATCTCTCAAGCAAATCTGATGACATGT
CCTACACACTCTCCTACTCTCCTACTCTGATAAAAGTCATAATTCTGGGATTAAATTCTAGCTGGTATTG
TCTGTGATTGGGCTGGTAAATTGTAACCTTATTGTTAGTACCCACATTTAACCTTAACTGAGGAAAAAATCTTC
AAGTAGACCTAGAAGAGGAGTTTAAAAACAAAACATGTAAGTAAAGGATATTCTGAATCTTAAATTCTACCCAT
GTGTGATCATAAACTCATAAAATATTAGTGTGAAAGGAGTACTTGTGTTAAATTAAAACACTCATGGATA
TGTAAGGAAACTGTCAGAATTTAATGTTGTTATTGTTATTGTAAGGAAATAGTGTGAAACAAAG
ATCTCTTCTACTGATACCTGGTTGATAATTATTGATGCAACAGTTCTGAGAATGATTCTGCAATTGCA
GAAATTAAAATCATCTGAGTGCAGGAAACATACAGTAAAGGAGGAGCAAAATACACATTGGGAA
AA

FIGURE 70

MGLFRGFVFLLVLCLLHQSNNTSFIKLNENNNGFEDIVIVIDPSVPEDEKIIEQIEMVTTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENKHADVIVAPPTLPGRDEPYTKQFTECGEKG
EYIHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSRACRIDSTKLYGKDCQFFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSED
FKNTIPMVTPPPP
VFSLLKISQRIVCLVLDKSGSMGGKDR
LNR
MNQAAKHFL
LQT
VENGSWGMVHF
STATIVNKLIQIKSS
DERNTLMAGLPTYPLGG
TSICSGIK
YAFQVIGELHSQ
LDGSEV
LLLTDGEDNT
ASSCIDEV
KQSGAIVH
FIALGRA
ADEAVI
EMS
KITGG
SHFYVSDEAQN
NGLIDAF
GALTSG
NTDLSQ
KSLQ
LESK
GTL
NSNA
WMND
TVIID
STVG
KDT
FFL
ITW
NSL
PPS
ISL
WDP
SGT
IMEN
FTV
DAT
SKM
AYLS
IPG
TAK
VGT
WAYNL
QAKAN
PET
LT
ITV
TSR
AAN
SSV
PPI
TVA
NMN
KDV
NSF
PSP
MIV
YAE
ILQ
GYV
PVL
GAN
VTA
FIES
QNG
HTE
VLE
LLD
NGA
GAD
SF
KND
GVY
SRY
FTAY
TEN
GRY
SLK
VRA
HGG
ANT
ARL
KLR
P
LN
RA
AY
IPG
WV
VNG
E
IEAN
PP
RPE
I
DED
T
QT
T
LED
FS
RT
AS
GG
AF
V
V
S
Q
V
P
S
L
P
D
Q
Y
P
P
S
Q
I
T
D
L
D
A
T
V
H
E
D
K
I
I
L
T
W
T
A
P
G
D
N
F
D
V
G
K
V
Q
R
Y
I
I
R
I
S
A
I
L
L
R
D
S
F
D
D
A
L
Q
V
N
T
T
D
L
S
P
K
E
A
N
K
E
S
F
A
F
K
P
E
N
I
S
E
E
N
A
T
H
I
F
I
A
I
K
S
I
D
K
S
N
L
T
S
K
V
S
N
I
A
Q
V
T
L
F
I
P
Q
A
N
P
D
D
I
D
P
T
P
T
P
T
P
D
K
S
H
N
S
G
V
N
I
S
T
L
V
L
S
V
I
G
S
V
I
V
N
F
I
L
S
T
T
I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCTGGAGTAGACTGACAGCAAAGACCGGGAAAGACCATACTGCCCCGGCAGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCCTATTCAAGGAAAGACGCCAAGTAATTTCACCCCA
GAGGAGCAATGATGTAGGCCACCTCTAACCTTCCCTTCTGAACCCCCAGTTATGCCAGGATTACTAGAGGTGTC
ACTCAACCAGCAAGCCTCCTCGGCTTAACITGTGGTGGAGAGAACCTTGTGGGCTGCCCTCTTAGCA
GTGCTCAGAAGTACTTGCCTGAGGGTGGACCAGAAGAAAGGAAAGGTCCCTCTGTGTTGGCTGACATCAGGAA
GGCTGTGATGGAATGAAGGTGAAACTTGGAGATTCACTCAGTCATTGCTCTGCCGCAAGATCATCCTTAA
AGTAGAGAAGCTGCTCTGTGTGGTTAECTCAAGAGGAGAACCTGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGCCCAAACCATGCTTCCTGTGCTAGGCCAGGGAAAGCCCTCCGTGGGGCCCGGCTTGAGGGATGCC
ACCGGTTCTGACGCACTGGCTGATCTGCTAATGATGATGTTGCCGGGCTGCTTGCGTGGATTTCCGGTGGTG
GTTTGCTGCTGCTCTCTGCTGTGCTATCTGCTGTACATGTTGGCCTGACCCCAAAGGTGACGAGGAGCAG
CTGGCACTGCCAGGCCAACGCCCCACGGGAAGGGGTAACAGGCCGCTTCAGAGTGGAGGAGCAGC
CGCAACTACGTGAGCAGCCTGAAGCGGAGATGCCACAGCTCAAGGAGGAGCTGCCAGGAGGAGTGGAGCAGCTCAGG
AATGGCAGTACCAAGCAGCGATGCTGGCTGGGCTGGAGACGGAGGCCCCAGAGAAAACCCAGGCCACCTC
CTGGCCTTCCTGCACTCGCAGGTGACAAGCAGGTGAATGCTGGCCTGAGCAGAGTATGCAAGCAGTG
CTTTTCGATGCTTACTCTACAGAAGGTCTACCAAGCTGGAGACTGCCCTAACCGCCACCCGAGGAGAACGCTGTG
AGGAAGGACAAGCGGGATGAGTTGGTGAAGGCCATTGAATCAGCCTGGAGACCCCTGAACAATCCTGAGAGAACAGC
CCCAATCACCGCTCTACACGCCCTGATTTCATAGAAGGATCTACCGAACAGAAAGGACAACGGGACATTGAT
GAGCTCACCTCAAAAGGGACCAACACGAAACAGCTCATCTTATTCGACCTTACAGCATTGAGCCCATCATGAAA
GTGAAAATGAAAGCTCAACATGGCCAACAGCCTTACATGTTATCGCTCTAGCAAAAAGGGTGGACAAGTT
CGGAGCTTATGAGATACTGGAGATGTGCAATTGAGCAGGATGGAGACTCCATCTACTGTTGTTACTTGGG
AAAGAAGAATAATGAAGTCAAGGAATACTTGAACACCTTCAAAAGCTGCAACTTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTCTGGGAAAGGGACTTGTGAGGCCGCTCTGGAAGGGAAAGCAACGTCCTTCTC
TTTCTGATGTTGACATCTACATCTGAACTCTCCTACATGCTGAGCTGCTGAATACAGCCAGGGAGAAG
GTATTATATCCAGTTCTTCTGCTAGTCACTCTGGCATAATATACGCCACCATGATGAGTCCCTCCCTGGAA
CAGCAGCTGCTATAAGAAGGAACTGGATTGGAGAGACTTGGATTGGATGACTGTCAGTATGGTCAAGAC
TTCATCAATATAGTGGGTTGATCTGGACATCAAAGGCTGGGGAGAGGATGTCACCTTATGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCCGTGGAGGACTCTCCACCTCTGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCCGAGCAGTACAAGTGTGCATGCGACTCAAGGCCATGAACGAGGATCCCACGCCAGCTGGCATGCTG
GTGTCAGGACAGAGATAGGGCTCACCTCGAACAGAAAGAACAGAAGTAGCAGGAAACATGAACTCCCGA
GAAGGATTGTTGGAGAGACTTTCTTCTTCTGCAATTACTGAAACTGGCTGCAACAGAGAAAAGACTTCCATAA
GGACGACAAAAGAATTGGACTGATGGGTAGAGATGAGAAAGCCTCGATTCTCTGTGTTGGCTTTTACAACAGA
AATCAAATCTCCGTTGCTGCAAAGTAACCCAGTTGAGCCTCTGAAAGTGTCTGACAAAGGCAAGATGTTGTG
AGATTATAAGCTAATGGTGGAGGTTGTGTTGTTACATACACTGAGACCTGTTGTGCTCATTGA
AATATTCTGATTTAAAGCAGTTGTTGAAAAAAATCATTGAGCAAGGCAACGATATTCTCTCATATGAATGA
GCCATCAGCAGGGCTAGTTCTAGGAATGCTAAATATCAGAAGGAGGAGGAGATGGCTTATTATGAACT
AGTGAATGATTAAGTAAATGGACGAAAAGAACATAAATATCGTGTATTTCCCAAGAT
TAACCAAAATACTCTTATCTTGTGTTGCTTTAACCTGTCCTGTTTCTTCTTATTTAAAGTCACT
TTTCTGCTTGTGAGTTAGTCTGCTTATTTAACCTCTGCAAGCCTTACAAGAGAGCACAAAGTGGCTAC
ATTTTATATTTTAAAGAAGATATTGAGATGCTTGTGAGGACTTCAAGTCAAAGCATCAAATTGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGCAAGGACTGAAATGTCAGGCAATTGAGACATAGGAGAAGGAAATGGTTGTACT
AATACAGACGTACAGATACTTCTGAAAGAGTATTGCAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
ACTTCTGCTTACAGAAAAGGAACCTCAGACTGGTATGTCAGTGTACCTAAAAGTCAGAAACACATTT
CTCCTCAGAAACTAGGGACGCCCTTCTACCTGTTAAATAACCAAAAGTACCCGTTGGAACCAAACAAATCTTTTC
AAAACAGGGTGTCTCTGGCTTCCATGAAAGAAGAATGGAGAAAATATATATATATATATTG
GAAAGATCAATCCATCTGCCAGAACTGAGTGGATGAAAGTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
TAACTGAATTATTTTAAATTAAGCAGTTACTCAATCACCAGATGCTCTGAAAATTGATTTTACCAT
CAAACATTTTAAATAACAGTTAACATAGAGTGGTTCTTCACTGAAAGTATTAGCCAGCACCAG
ATGCATGAGCTAATTATCTCTTGTGAGTCCTGCTTGTGCTACAGTAAACTCATTGTTAAAAGCTTCAAGAAC
ATTCAAGCTGTTGTGTTAAAATGCAATTGATGTTACTGTTACTGAGATTATGAAATTAAATTAACACAGG
CCATGAATGAAAGGTGTTGCAAGCTAATAAAATGATTTGTGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVKNKEKLNMAN
TLINIVPLAKRVDKFRQFMQNREMCIEQDGRVHLTVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIIYGHHDAPPLEQQQLVIKKETGFWRDFGFGMTQCQYRSDFINIGGFDLDDIKGWGGEDVHLYR
KYLHSNLIVVVRTPVRLFHLWHEKRCMDELTPEQYKMCMQSKAMNEASHGQLGMLVFRHETEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGAGGGCAAAGAGGCAGCAAGAGAGTTGTCCTGGGATCCA
GAAACCCATGATAACCTACTGAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCTCACTCCTC
CCTCCCTCTCTGCCTGCTCTAGTCCTCTAGTCCTCAAATTCCAGTCCCTGCACCCCTC
CTGGGACACTATGTTGTTCTCGCCCTCCTGCTGGAGGTGATTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACATTGCCAGCCTTACCCCTGAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCGTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGGCC
ACACAGTGCAACTCTCTGCCCTCACCCCTGTATCTGGGTGGACTTCCCCGAAAATATGTAGCT
GCCAGCTCCACCTGCACTGGGGTCAGAAAGGATCCCCAGGGGGTCAGAACACCAGATCAACAG
TGAAGCCACATTGCAAGAGCTCCACATTGTACATTATGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGGCTCAGGGCTGGCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTGAGTCATTGCAAGTCAGGCTAAAGATCAGAACAGACCTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGTCACAACCTCCCCCTGCTACCAAGAGTGTGCTGGACAGTTTTTATAGAACAGTCCCAG
ATTTCATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCCACAGAACAGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTGCTTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTCTGCCCTCTGGCTTTATTCTAGAAAGATTGGAAGAACAGGAGGCTGGAAAA
CCGAAAGAGTGTGGTCTTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCCCTCTCAGATAC
CATGGATGTGGATGACTTCCCTCATGCCATCAGGAAGCCTCTAAAGGGTGTAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTTCAGCCTCTCAAACATGTAGGAGGA
ATGAGGAATCGCTGTGTTAATGCAAGAGANAAACTCTGTTAGTTGCAGGGGAAGTTGGG
ATATACCCCAAAGTCCCTCACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCCTAGGATAAAGAGTTGCTGTTGAAGTTGTATTTTGATCAATATAATTGGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECNNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLOQTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCCCTGGGGACGGGCAGTCCCTGTGTC
TCTGGTGGTTGCCTAACCTGCAAACATCACCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTGACAGCTCC
AGAGAAGTGGAAAGAGAAATCCAGAAGACCTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCTCGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTTCCTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTTATGAAATGAATTGACAAAGATTCTTGTGCCTGCTGAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTCGGATGATTCTAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCAGGGAAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGGTATGCTTCGCAATTGATGAAATTGACTCTGAAG
AAAACACCGAAGGTACTTCTCACCCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTGGGCCCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTGATTCCCTCGCTGTCAGCTCGACCAGGATTCAAGGGCTGCG
AGCCTCTGAGGGGATGGCTGGAGAGGGTCTCTATCTAGACTCTAGGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCTAGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGCCAACACTTCCTTGCCTTGTGCTGAAACAAGTGAG
TCACCCCTTGATCCCAGCCATAAAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCAGTGT
CTGTGAGAATTACTTATTCTTTCTTCTATTCTCATAGCACGTGTGATTGGTCATGCATGTA
GGTCTCTAACATGATGGGGCCTGGAGTCCAGGGCTGCCGGTTGTTCTATGCAGAGAA
AGCAGTCATAATGTTGCCAGACTGGGTGAGAATTATTACAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSLNTKSNRWSQCVTNHTLVLWLEPNLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGYNEFDKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPQQEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQQLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDDGLGEEGLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQmen

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTTGACAC
CTGGGAAGATGGCCGGCCCGTGGACCTTCACCTCTGTGGTTGCTGGCAGCCACCTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCACTGCAGCAGCTGCCGCTGCTCAGTGCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCAAGATCAGCTAACATCCTCCAGCTGCAGGTGAAGCCTCGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAACGCACCATCCGCATGGACACCAGTGAAGTGGC
CCCACCCGCTGGCCTCAGTGACTIONGCCACCAGCCATGGGAGCCTGCGCATCCAACGTGTA
TAAGCTCCTCCCTGGTGAACGCCCTAGCTAACAGGTCAAGCAGCTCATGAACCTCTAGTGCCATCCCTGC
CCAATCTAGTGAACACCCAGCTGTGTCCCGTATCGAGGCTTCCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCATTCAAGCTCTACCTGGGGCCAAGTGTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGTGCTCTCCAGAAGA
ATTCAATGGCCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAGCATGGC
TGATCAATGAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTGAAGATCTAACTCAGGAC
ACTCCCGAGTTTTATAGACCAAGGCCATGCCAAGGTGCCAACTGATCGTGTGGAAGTGT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTACCCCTGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACAGAAATGGCAAATTAAAGATCTGGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTCGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGAAAGGCTGG
GTCCCAGCTGGAGTATGGGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTOELKDHNATSILQQPLLLSAMREK
PAGGIPVLGSLVNTVLKHIWLVKVTANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMPMTLDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLGIEASSEAQFYT
KGDQILNLNNISSLRQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTGTGGCTACATCTAGGCCCTCTGGGCTTTGGGCACACTGGTGCCT
GCTGCTCCCAGCTGGAAAACAAGTTCTATGCGGTGCCAGCATTGTGACAGCAGTTGGCTCT
CCAAGGGCCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTCTGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGCACATCCAGTGCAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCAAAGACAGAGTGGCGTAGCAGGTGGAGTCTTTTATCCTTGGAGGCCCTGGGATT
ATTCCCTGTTGCCCTGGAATCTCATGGGATCCTACGGACTTCTACTCACCAGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTTTACTGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTCCCTGCTCATCCAGAGAAATCGCTCCAACACTACGATGCC
CAAGCCCAACCTCTGCCACAAGGAGCTCCAAGGCCGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCAAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTGAAAACAGTGGACAGCACCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTTCC
TCACCTGCTGCTCCCTGCCCTAAGCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGGCTCTGGTTACCTGGACTCCATCCCCAAACCCACTAATCACA
TCCCACGTGACTGACCCCTCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTCTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTGTGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCTGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTIVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSIASSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTCGTCTCTCCATCTCTCCCTCCTTC
CCCGCGTTCTTTCCACCTTCTCTTCTCCACCTAGACCTCCCTCGCCCTCGCTAGCAGCAGACCTCCCTGGGCTGTGG
GCCACCGCTGCTTCCTGGCCCTCTCCGACCCCGCTCTAGCAGCAGACCTCCCTGGGCTGTGG
GTTGATCTGTGGCCCTGTGCCTCCGTGCTCTTCTGCTCCCTCCCTCCGACTCCGCTCCCG
ACCAGCGGCCTGACCTGGGAAAGGATGTTCCGAGGTTCCTCTGCTGG
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGGCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCACCAAAGTCCCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCTGCCAACCAAGTGTGCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCACTGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCGGGACCCCCAGCCCCACTGGCCTCAGGCCCTCTGAGCTCATCCCTGCCACTT
CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGCCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCACTGGCGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCTTCCGTGCCTCGGC
CCCTGCCCTGCATCCTATGCCACTGTGAGGATGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCCTGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCAAGCCAGACAACCTGCGTCGCTTGCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGCCAAGGCCACACAGCCAGAATCTTCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCAGTCCACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTGCAGATGTAATAA
TAAGAAGTTGCATTACCTCAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTTCPPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGBTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKKKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTLPTARWPPRRSLERLPSPDPAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCCGCTACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTCTCCTCTCTCTAATCCATCCGTACCTCTCCGTCA
TCCGTTCCATGCCGTGAGGTCAATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGGTGGTGGCAGAACAGCCTGTCCAGGCCCTGGTGGGGAG
GACGCAGCATTCTCTGTTCTGTCTCCTAAAGACCAATGCAGAGGCCATGGAAGTGCAGGTTCTTCAGGGG
CCAGTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGGCGATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGCCCTCATGGTGCAGGATTAGTCCCAGTCTTACTACCGAACGCCATCTGGAGCT
ACAGGTGTCAGCACTGGCTCAGTTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCCGGCCACAGCGAAGTGGAAAGGTCACAAGGACAGGATTGTC
ACAGACTCCAGGACAAAACAGAGACATGCATGCCCTGTTGATGTTGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATCTGTTCCATGCCGTGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACTTTTCGAGCTATATCGTGGCACCTGGCTACCAAAGTACTGGAATACTCTGCTGTC
TTTTTGGCATTGTTGACTGAAGATTTCTCTCCAAATTCCAGTGGAAATCCAGGCGGAACGGACT
GAGAAGAAAGCAGGACAGGCAGAATTGAGAGACGCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTCACCGAAGCTCTGCGTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGACTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGAGGTGGACGGAGGACACAATAAAAGGTGGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCCGATCATGGTACTGGTCCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGTCTCCCCAGGACCCACCTACAAAATAGGGTCTCTGGACTA
TGAGTGTGGGACCATCTCCTCTCAACATAATGACCAGTCCCTTATTATACCCCTGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAATGGAACCTCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAACAGAGGCCCTTGGCAAAGGGCCTGCAATCCCAGAGAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTTCCCTCCCCAGGGTCAAATGTAGGATGAATCACATCCCACAT
TCTTCTTCTAGGGATATTAAGGTCTCTCCAGATCCAAGTCCCAGCAGCAGCCGGCAAGGTGGCTCCA
GATGAAGGGGACTGGCCTGTCACATGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGGCTAAGTGTGATCTTGAATACCAACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGTGGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTCATTATATTACACTTCAAGTAAAAAAA

FIGURE 84

MAIMLSLVSLKLGSQWQVFGPDKPQALVGEDAASFCLSPKTNAEAMEVRFFRGQFSSVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPOQQDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEILDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPRTPPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTTAACCCAGACATGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGGGCGGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTACCCCTCGCATGGCTGGATT
TACCTGGCCCAGTAGTTCATGGCTACTGGTCCGGAAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTGGCAGTGTGGAGGAGACTGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGTGATGCCGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGAAATTATAAACATCACCGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGCCTGTGAGCAGGGACACCCCTATGATCTCTGGATA
GGGACCTCCGTGCCCCCTGGACCCCTCCACCACCCGCTCTCGGTGCTCACCTCATCCCACA
GCCCGAGGACCATGGCACCAGCCTCACCTGTCAAGGTGACCTTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTAACGTGCTTACCCGCTCAGAACTGACCATGACTGTCTCCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTCATCTGTCACTCCCAGAGGGCCAGTCT
GCGCCTGGTCTGTGCAAGTTGATGCAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCCTGACCCCTGTGCCCTCACAGCCCTCAAACCCGGGGTGTGGAGCTGCCCTGGGTGAC
CTGAGGGATGCAAGTGAATTCACCTGCAAGAGCTCAGAACCCCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAAGACAGCAACATCAGGAGTGAACAGCAATCCCCCTGCCAGGCTGAGCTGGAG
CCACAGCCCTGGTCTTCCCTGCTCATCTCGGTATCTCGTTGTAGTGAGGTCCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCTGGAGATAAGGGCATAGAGGATGCAAACGCTGTCAAGGGTTC
AGCCTCTCAGGGGCCCTGACTGAACCTGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCGCTCTCAGTGGGAAAGGAGAGCTCAGTATGCATCCCTCAGCTCCAGATGGTG
AAGCCTGGGACTCGCGGGGACAGGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCGTATTGAGGGATCACAGCCCTCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAGAACACAGGCTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAGACAGACAAATTCTA

FIGURE 86

MLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVNTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPDHGTSLTCQVTFFPGASVTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCAVDADSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAEFTCRAQNP
LGSQQVYLNVSLSQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSAQGPLTEPWAEDSPPDQPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAAGCTGCAAGGAATCAAAGACGAATGTCTAGTGCATTG
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGT
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTGGATCTGCAGAGGCCACGGCAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAATATGGAGAAGGAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCAGCAGCCAGAAACAGCATCTTATTACTCACCCAT
GGCCAGCGGAATTCACTGCGGATTGTTAGTTCACTGAGGTATTAAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTACCGGATGTAACACTGAGCATCACTGCATTGGTGAGGAG
GATACTTCCAGAGGCCAGTCCCAGCAGTGTGGAGATTTCTGGTTTGATTGGAGTGGATAT
GGAACCTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAAACCCAGACCTCTCCCAACCATGAGATCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAAAT
SDDYKNPGYYDIQAKDLGIWHVPNPKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVIKVYDFGDAQKTASYYSPYGQREFTAGFVQFRVFNNERAANALCAGMVRTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTGTCCCCGAGGCCATGGGCCGGTCTCAGGGCTTGTGCCCTCGCTTCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCAGGACAGAACATACAGGCCTG
CCTGCCTCTCACGTTCACCCCCGAGGAGTATGACAAGCAGGACATTAGCTGGTGGCCGGCTCT
CTGTCACCCCTGGCCTTTGCAGTGGAGCTGGCCGGTTCCCTCTCAGGAGTCTCCATGTTAAC
AGCACCCAGAGCCTATCTCATTGGGCTACTGTAGTGCATCCGTGGCCCTGTCTTCCAT
ATTAGCAGCGTTGGAGTGCACTACGTATTGGTACATTGGTCTCTGCAGTGCCTTCCAGCTG
TCACTGAAATGGCTTATTCGTCACCGTCTTGGGCTGAAAAAGAAACCCCTCTTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGGCGCTCGTATTCTGGAAAGAAGGAAG
GCATAGGCTTCGGTTTCCCTCGGAAACTGCTCTGCTGGAGGATATGTGTGGAATAATTACG
TCTTGAGTCTGGATTATCCGATTGTATTAGTGCCTTGTAAATAAAATGTTTGTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKP

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAAGATGCACTGAC
TCGCTGCTGTTCTGTTCTGGTCAGGGTAGCCTCTATCTGGCATCTGGCCAGGATGATG
GTCCTCCCGCTCAGAGGACCTGAGCGTGTGACCACGAGGGCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGACACATCTCACCTAACGTCGGCCCATGGCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCCCTAGCCAAGGTGAAGAAAATCTTGGCTGGGCACCTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCATCAGCCTCGTGCCTCCAGTAAAGCTGTAG
AGTCCACCAAGGAACAGCAGATCTCATCGAAGCCAAGGCCTCCAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
TCTACATCGCCTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACATAC
CATAGTGTACCCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAAGGACAGGCCCTGCCATGCAGGAGACCATCTGGACACCCGGCAGGGAAAGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCAAACCTGAAGCTGTGGAGTGAATGACTCACAGGAGCACTGG
AGGAGGGAGTGGCTCTGTGTCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCGAGGCCCTGTGGCAGGCCATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGGAGCTCCAGCCCTGCCTAGTGGCGCCCTGAGCCCCCTGTGCTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCCTGGGGCTTGTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTTGTGCCAGTACTCCCCCTGTACCAACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACAGACAGGAGATTGTGGCTCCACACTAACGGCCACAGCCCAC
CGCGTGCTGTGTGTCCTCTTCCACCCCAACCCCTGCTGGCTCCCTGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTCAAGACCGGGGTTCTCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGTGGGCGGGCCGGCGCAGAGCATGTGCTGGATCTGTC
TGTGTGCTGTGTGGGGGGGGAGGGAGGGAGTCTTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTGTTCTGGAGCAGGAATAAGCTTGGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILVVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTCGGCTGCCCTCATTGCCCTCGGGCTGCGCTGCC
CTTTATGTCTTACCATGCCATCGAGCCGTTGCGTATCATCTTCTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCAATTATAAACTCTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCATGCGACTGCTGGCCTATGTTCTGGCTTGGGCTTGGAAATCATGA
GTGGAGTATTTCCCTTGTAATACCCATCTGACTCCTGGGCCAGGCACAGTGGGATTCA
GGAGATTCTCCTCAATTCTCCTTATTCACTGACTTGTGAGCCTGGTCATTATCTTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTATCGTTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTTCATAAGTTCTATTATGAAATAAACCTGGCG
TCAGCATTATAATCCTGGTGTCAAGGCCACTGGCATTCTAGCTGCGGGAGGCAGCTGCC
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACCTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAAGCACTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAGAAATAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAFGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISLIVWFMARVIIDNKG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPCTVGIHGDSPQFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLTL
HLLVSAQTFISSLVYGINLASAFIILVLMGTWAAGGCSRSLKLCLLCQDKNEFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGGAGTACGATTCTACTGTTTGCTTCTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAAACCAACAGCAGTCAAATCAGGTCTTCCTTCTTAAGTCTGATACCATAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTCACACAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCTGTCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCCAGCAGGCCGCTCCCAACTCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTCAAGAAAATATGTTATGTAGAGAATCCAACTTTAAAAAA
CAATAATTCAATGGATAAAATCTGCTTCAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTIGPDLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEQQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWSLEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDYAAPLEHVL
VMDENDNVPICPYRDPPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNI LLLVIA MDLAGAEGGF SSTCEVEVAVTDINDHAFEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTEGT FGLDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVLGPGPGP GATATVTVLVERVMPPP KLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTIAPVPS
QYLCTPRQDHGLIVSGPSKDPD LASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTIAHWEP
REHIIIPVVVSHNAQM WQLLRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLI
LIFT HWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCAACGCCTG
AGTCCAAGATTCTCCAGGAACACAAACGTAGGAGACCCACGCCTCTGGAAAGCACCAGCCTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAATCCTCTGTTCTTGCCTCTAAAGTCTGGTAC
ATCTAGGACCCAGGCATCTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCCTTATGTTGGTCTACTATTGCATTAGAAGCTGCAACAAATTCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACGCCACCATCTCAGGGTCCAGCGTACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGGCAGCACAGTCACCAACT
CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCAGTGGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCAGTGGGGCAGCACAGCCAC
CAACTCTGACTCCAGCACAAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
GCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGGCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCAGCCTCCAGTGGGCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCAGCACTGCCACCAACTCTGAG
TCCAGCACGCCACCTCCAGTGGGGCCAGCACGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACGCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGGCCGGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAAGCACGCCACCAACTCTGAGTCCAGCACAAACCTCCAGTGGGGCTAGCACA
GCCACCAACTCTGACTCCAGCACAAACCTCCAGTGGGGCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACAAACCTCCAGTGGGGCCA
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAAGGCTCTGGAACAGCAGCTGTGACTGGA
ATGCACACAACCTCCCATAGTCATCTACTGCACTGAGTGAGGAAAGCCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTCCTCATCACCCCTGGTCTGGTTGTGGCGGGCGTGGGGCTCTTGCTGGC
TCTTCTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTAAACACAGCTGTACCCACCT
CATGCCCTCAACCATGGCCTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAACTGGTCTGGAGGAGACCAAGTATCATCGATACCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTGAGCAGCCCGGAAGCAAGTGCCTGCATTCTCAGGAAGGAAGAGACCTGGGCA
CCCAAGACCTGGTTCTTCATTCATCCCAGGAGACCCCTCCAGCTTGTGAGATCCTGAA
AATCTGAAGAAGGTATTCCCTCACCTTCTGCTTACAGACACTGGAAAGAGAAATACTATAT
TGCTCATTAGCTAAGAAATAATACATCTCATCTAACACACAGCAAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACACTGAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMKGKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNKGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGENHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCGCGTTACGGGATGAATTAAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCACCGCACCCACTCGCGTCGCGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCCTCTGAACCCATGGTCATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTCACTGGAAGGGATTACTTACTGACTCTGTTGGGAAGCTTTGGAAAGCATTT
TCATGCTGAGTCCTTTTACCTTGATGTTGAAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTACCCCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAAGT
GATTATAACTGGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCATGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTTGGGCATGCGAGGCTGCTGCCTATAT
CTTCATTCACTAGGAAATGGAAGGATGACAAGAGCCATTGAGACATGATTGATTACTTTGTG
ATATTCACTGAACCACTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCAGGAGACTTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCCTCCCATCCAAGGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGGCTGCGTTCTTCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCCTGTTAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTATAATCACCATTGAAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACATTGCACTGTTACCGACTTTACACAAACAGCCACAT
TTAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAACCTAGAGCATATTG
GAAATGTTCTAAACCTTCTAACGCTCAGATGCATTGCACTGACTATGTGAAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGCAACTGTGACATAGCAGGGAGTGTGATGGGGTGAATAACTT
GGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLERKICLKASLGVPFGWAMQAAAYIFIHRWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFPREIHFHVHRYPIDTLPTSKEQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCCACGGCTCTGCCCTGAGACAGCTGGCTGACC
TCCAAATCATCCATCCACCCCTGCTGTATCTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTTCTCAGTTCTACGAGCTGGTGTCAAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCCTGGTGGGGGAGGACGCCGTGTTCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGGTCTTCAGGAATCAGTCCATGCTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTAAGGACTCCATTGCAAGGGGGGCTGTCCTCTAAGGCTAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTCAGTCCCAAGTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGCTCATTCCTCTCATTCACCTGTTGGGATATGTCAGGGAGGTATCCA
GTTACTCTGCCCTGCCCTCAGGCTGGTCCCGACAGCCACAGCCAACTGGAAAGGTCACAAAGGAC
AGGATTGTCTCAGACTCCAGAGCAATGCAAGATGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTTGGCCTGGCTCTA
TTTACTCGGGTTACTCTGGTGCCTGTGTTGTATGGGATGATAATTGTTCTTC
AAATCCAAGGAAAATCCAGCGGAACCGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGGTGTGGTGGCTCTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGTGGTGGTATGTGGAGTGTGTCGGGATGACGTAGACAGGGGAAAGAAC
ATGTGACTTTGTCCTCCAACAATGGGTATTGGTCTCAGACTGACAACAGAACATTGTT
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACACGAGTAGGGTCTT
GGACTATGAGGGTGGGACCATCTCTTCAATACAAATGACAGTCCTTATTATACCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATTATCATATGTCAGTGCCTGGGATTAGACAGAGAACAGGCTGTTAAAGGGC
CCACACCACAGACCCAGACACAGCAAGGGAGAGTGCCTCCAGGTGGCCAGCTT
CCGGAGCCTGCGCACAGAGAGTCACGCCCTACTCTCTTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCAGCGCAGTCACAGCTCCAGATGAGGGGGATTGCCCTGACCTGTGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGAGCGAATAGACTCACATTAGGTTAGTTGTGAAAA
CTCCATCCAGCTAACGATCTGAACAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAGAGACAAACGAATGTGAATCATGCTTGAGGTT
TGAGGGCACAGTGTGCTAATGATGTGTTTATATTACATTTCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACAAATCACCATGGAATAGTTATTGAACACC
TGCTTGAGGCTAAAGAATAAGAGGGAGTAGGATTTCACTGATTCTATAAGCCAGCAT
TACCTGATACCAACAGGAAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAAATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTTGGTTAATATTAAAT
ATCAACCAGTGAATTCAACATTAATAAGTAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCSSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCALCGVVMGMIIIVFFK
SKGKIQAELDWRRKHGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGVCRDDVRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCAATGATGATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGAAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTACAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGGCTCTAACATTACAGAAATGAGCCAGAGACTGAAT
CAATGGTAAAAATGCATTTATAAATCTCCATTAAAGGAAGAATTGTCAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTCACATTGTTTACATGAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAAACCTAGGTCAAGAGTCTCAGGATCAGGATGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGAGTCATCGCTGTGGAGCAACCTTAATTATGCCACATGGC
TTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCCTGCCAGATGGACTGCTCCCTGGAGTAACA
ATAAAACCTCGAAAATGAAACGGGGTCTCCGAGAAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGAGAGCTTCTAGCCCTGTCACAAATGCAGTACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAAT
GATGGTTACAGTCAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTAGCTAGAGATACTGGTACCTTGCTGGAAATAGTG
AGCTGGGAGATGAATGTGCGAACCCAAACAAGCCTGGTAACTAGAGTTACGGCCTTGCAGGGACTG
GATTACTCAAAACTGGTATCTAAAGAGACAAAGCCTCATGGAACAGATAACATTTTTTGTTTTG
GGTGTGGAGGCCATTAGAGATAACAGAAATTGAGAAGACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGCTTGATGCATGTATTCTCCAGCTCTGTTCCGCACGTAAGCATTGCTGCTGC
GATCAACTCTGTCATCTGTGAGCAATAGTGAAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTGTTCTAGAAGTTGTCAGAATTGACTTGTGACATAATTGTAAT
GCATATATAACATTGAGCAACTCCTTTCTTCAGTCCCTCAGCTCCTCTCATTTCAGCAAATATCCATT
TCAAGGTGCAGAACAGGAGTGAAGAAAATAAGAAGAAAAATCCCTACATTATTGGCACAGAA
AAAGTATTAGGTGTTCTTAGGAAATTAGAAATGATCATATTCAATTGAAAGGTCAACCAAGACA
GCAGAACACCAATCACTTCATATTAGGAAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGGCCAG
ATATATCCTTATTTCATTCCAAACAACTACTATGATAATGTGAAGAAGATTCTGTTTGACCT
ATAATAATTATAACAAACTTCATGCAATGTACTTGTCTAACAAATTAAAGCAAATTATTAACATTG
TTACTGAGGATGTCACATATAACAATAAAATATAACACCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVI FISLIVLAVCIGLT VHYVRYNQKKTNYY STLSFTDKLY
AEFGREASNNFTEMSRVKNAFYKSPLREFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLH EKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHP SHDYDISLAE LSSPVPYTN AHRVCLPDAS YEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATT CNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDI WYLAG
IVSWGDEC A KPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTCCCTG
CCCCGATGAGCCCCCGCCGTGCGTCCCCACTATCCCCAGGCAGGCGTGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTTGCCTTGGGAGTAGGATGTGGTAAAGGATGGGGCTTCTCCCT
ACGGGGCTCACAAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCCTGCCTACGCCCTCAA
TCTGCTTTGGTTAATGTCATCAGTGTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTACTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTCT
GTGGTCATCCGGTATGATTGCTGTTGCTGTTCTTATCATGTTGGGGATGTTAGGATATTG
TGGAACGGTGAAGAAGAAATCTGTTGCTCTTGACATGTTGGAAAGTTGCTTGTCAATTTCT
GTGAGAACTGGCTTGTGGCTTGACATGAAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGCAGGATGACAATTATGGATTACCTAGATATCGGTGGCTACTCA
TGCTTGAATTTTTTCAAGAGAGAGTTAAGTGTGTTGGAGTAGTATTTCACTGACTGGTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCCTGCTGTTAGAGAATTCCAGGGATGTTCAA
CAGGCCAACAGGAAGATCTCAGTGACCTTATCAAGGGGTTGGGAAGAAAATGTTCC
TTGAGAGGAACAAACAACTGCAGGTGCTGAGGTTCTGGGAATCTCCATTGGGTGACACAAA
TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCTGG
ACAGACCAATGATGTCCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCTCAGTAGAACT
GTTGAAACCAAGCCTGTCAGAATCTTGAACACACATCCATGGCAAACAGCTTAAACACACT
TTGAGATGGAGGAGTTAAAAAGAAATGTCACAGAAGAAAACACAAACTTGTTTTATTGGACT
TGTGAATTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAATGTTGCCATAAAA
TAACACCTAAGCATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTCATACCCACTGTCAG
TGTGATGACTTTACTGAACACAGTTATGTTGAGGGCAGCATGGTTGATTGCAATTCCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGAGCCATAGTAAAGGTTGATTTACTCTACCAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATT
CTCAGCGATCTATTCTCTGATGCTAAATAATTATGATGAAACTTCAATAATTGGTACT
ACCTAAATGTGATTTGCTGGTTACTAAAATATTCTACCACTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGACAGGATTTGATATAAGCTGTGTTAAATCTGTATAATTCACTGATCGAT
TTCAGTTCTGATTAAGTAAAGATAACCAATTGAAAGAAATTGTCCTGTATAGCATCATT
ATTTCAGCTTCTGTTAATAAGCTTACTATTCTGCTCTGGCTTATATTACACATATAAC
TGTATTAATACTTAACACTAATTGAAATTACCACTGTCAGTGTGATACATAGGAATTCATT
AGAATGTAGTCTGGTCTTAGGAAGTTAATAAGAAAATTGCAACATAACTTAGTTGATT
AAGGACTTGTATGCTGTTTCTCCAAATGAAGACTCTTTGACACTAAACACTTTTAAAAA
GCTTATCTTGCTTCTCCAAACAGCAATAGCTCCAAGTCAATATAAAATTCTACAGAAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACACAGATTAAATTGTTACAA
CTTGTGTTATTCTACTGATTAATATACTGTGGCAAATTACACAGATTAAATTGTTACAA
GAGTATGATTAATTGAAATGGGAAAAGTGCATTACTGTTACTGTTACTGTTGTTAT
TTCTCAGAATATGGAAAGAAAATTAAATGTCATAAAATATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTLLTAETRVEEAVILTYFPVVP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVYFTDWLEMTEMWDPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGGKMKYSFLRGTKQLQVLRLGISISGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACCTCATCCTCATCCTCTTCCTCTGATAAAGCCCTACCAAGTGCT
GATAAAGTCTTCTCGTGAGAGCCTAGAGGCCCTAAAAAAAAAAAGTCTTCTGGCAGTTGGTCCAGAAGGGATGCCTCCATTCTGCTCACCTG
CCAGTATTAAGAGGATTTCCAGTGTTCAGTGGCAGTTGGTCCAGAAGGGATGCCTCCATTCTGCTCACCTG
CCTCTTCTACAGGCACCTCCGTGTCACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCACGGGCATGGGGGAGATGCCATGCCTACCTCTGCATACCAGAAAACACTGTGGAACCC
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAACGGCAGGCATTGTGCAACGCCAGGTTGTGCCAGCT
TCAATGGGAACACTGCTGTCTGGAACACCACGGTGGAGTCAGGCTTGCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGTCATTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCAGTGCACATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGCACTGAGATCTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGGTGGCGTGTGCTAAGAAGTGTGCAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTGGCTTGGATCTGAGAAAGGCTACAGTGTGAAATGTCCCCGGGCTGGTGCTGT
CTGAGGATAACCACACTTGCACGTCCTGTGTTGTGCAATCAAAATGCCATTGAAAGTGAACATCCCCAGGG
CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCTGCCAGGAGTGTCAACGCCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTGATGTGGTGAATGACAAGATTGTGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCCCGGGAGCAGGGGACTTCATCATCCGAAACCAGCAAGCTGCTGATCCGGT
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAAACCTTCGAAACACTCCCCACTGGAAAT
CATGAGCCGAATCATGGGATCTTCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTCGTACCTCTACTTGGCATTGAGCCGTGGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTGCACCCACCTCCAAGATGACGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAAGATGACTCGGTAAGCAGTACACATCCGGGATCACCTAGCAAACCACTTCCAGG
TCCCTGTTCAAGTTGTGGCAAAGACCAAGGAAGTGTGACTCCCTACTTGGCATTGAGCCGTGGCACGTGAGTG
TTGGACGGCGTCCCGCTGTGCCACCCACCTCCAAGATGCGTCTGGGAGGAGAGGACTCAGC
CGGTCTACAGGGCAGACGCTAACAGGGGCCGATCCGCATCGACTGGAGGGATTCGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTTGGAGCTCTCCCCCACCGCCCTCAAGAACATCTGCCAACAGC
TGGGTTCAGACTTCACACTGTGACTTCAGACTCCACGCCAACTCACTGATTCTGCTTCCATTCACTGGGCA
CAGGTACAGCAGCTGAACAAATGCGCTGGGTTTCATCTTCTAGGGTTGAAAACACTAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTACCTAAATACCTCGTGTATGGTGCACATCAGAC
CACAAAATCAGAACGCTGGGTATAATTCAAGTTACAAACCCCTAGAAAATTAACAGTTACTGAAATTATGA
CTTAAATACCCAATGACTCCTTAAATATGTAATTAGTTACCTTGAATTCAAATGCAACTAA
TTATAGGAAATTGGAAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSRGKTCEDVEGCHNNNGGSHCLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFAATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCCTCCAGCCAGTGACCAAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCC
CCTGCTGCCTTGGGTGACAATCTCAGCTCAGGCTACAGGGAGACCGGGAGGATCACAGGCCAGC**ATGT**
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCAGTATCATCATTGT
GGTTGTCCTCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCTTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTCTCTGCCTGTTGACAACCTCACAGAGCTCGTGAAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGATGTTGTAATCACAGAAAACAGCCAG
GAGCTTCGCACTGGGAACTCAAGTGGCCCTGTCTCAGGCTCCCTGGCTCCCTGCACTGTCTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGTGGGGAGGAGGCTCTGTTGATTCTGGCTTGGCAGG
TCAGCATCCAGTACGACAACAGCACGTCTGTGGAGGGAGCATCTGGACCCCCACTGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCTGGCTGTGGCAAGATCATCATCATTGAATTCAACCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTCAGGCACAGTCAGGCCATCTGCTGCCCTTCTT
GATGAGGAGCTACTCCAGGCCACCCACTCTGGATCATTGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGCTGACATACTGCTGCAGGCGTCAGTCAGGATCATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGAAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACAGTGGCATGTGGTGGCATCGTTAGCTGGCTATGG
CTGGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCTAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCT**TAAT**GCTGCTGCCCTTGCAGTGCAGTGGAGCCCTTCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTGGGTACACCCCTGCCCACGCCAGCAT
TTCTGGAGCAGCAAAGGGCTCAATTCTGTAAGAGACCCCTCGCAGGCCAGAGGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGCTCCAGCATCCAGGGAGAGACAGCCACTGAACAAGGTCT
CAGGGTATTGCTAAGCCAAGAAGGAACATTCCCACACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTCACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIALLSIASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGEELDCPLGEDEEHCVKSFPEGPAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAAETACRQMGYSRAVEIGPDQDLDVVEITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGGEEASVDSWPWQVSIQYDKQHVCAGSILDPHWVLTAACFRKHTDVFNWVKVRAGSDKL
GSFPLSLAVAKIIEFNPMYPKDNDIALMQLQFPLTSGTVRPICLPPFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTQGDGGPLMYQS
DQWHVVGVIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACATTTCAGCAACTAAAAAGCCACAGGAGT
TGAACGTCTAGGATTCTGACTATGCTGTGGCTAGTGCTCTACTCCTACACTACATTAAATC
TGTAAAAACGGTCTCTTGTAACTAGCCTTACCTTCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCGCTGGAGCAGTGCCCTACCAACTGTCACGTCT
GGAGGCAGTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCTGGCGTAGAAGGGATTTGACAAGCCGAAGATTCTAGGCGATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGCAATCACTGCCCTGGGCCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAACTGAACCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGTCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAGCAAAGTGGCAGGGAAAGGAACCTGTGCCAAATTATGGGTAGAAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCTGCATTCACTGGACATGTGGGGAAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCACGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGCTGGAAAGGAGCAAGCAAAGTGACCAATTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCTACTAAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTAGCAAGGCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGAAATGCCCTCAGTAAGCACAGGCTGAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTAGTAATTGGAGCTGTCCCCAAGACCAAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGGCCCTTATACCCAGGAGACTTGATTTGAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGCTGTAATGCCAACAT
TTTGGGAGGCCAGGGCGGGTAGATCACCTGAGGTAGCAAGACCAAGGCCATGGCCAAACATGG
TGAAACCCCTGCTCTACTAAAAACAAAAAAACTAGCCAGGCATGGGGAGGTGAAGGAGGCTGAGACA
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTGAAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAACATTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLLPTLKVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAACCTCCAGAGAATTGTGAAGACTGTCACATTCAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTGGGGAGCAAGCACTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGAAATTGATCCTGTGACAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTAAACGGATAACTGG
CATCTACTCGTGGTCTTCAAAATCTTATCAAACACTCAGATTAAGTGAATTCTGAATT
CTGAACCCAGAAGAGGAAATAGATGAGAATGAAAGAAATTACCAACTTTCTTGAAACAGTCAGT
ATTTGGTCCCAGCAGAAAGCCTATTGAAAACCGAGATTTCTAAAAATTCCAAAATTCTGGA
GATTGTGATAACGTGACCATGTATTGGATCAATCCACTCTAATATCAGTTCTGAGTTACAAG
ACTTGAGGAGGGAGGAGAAGATCTCCTTCCCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCAATAATGACTATACTGAAATGAAATAGAATTGATCCATGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGACGAGTCATCTGCGTCATCATGCCCTG
TAACTGGTGGTGGCCCGCATGCTGGGAGGGTTAATAGGAGGTTGAGCTAAATGCTAAAC
TGCTGGCAACATATAATAATGATGCTATTCAATGAAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCCAGAATTACTGTAGGTATTCTCTTCTATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTILVLFWGSKHFPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEETTTFFEQSVIWWPAEKPIENRDFLKNSKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGCAGCAGGAGGGCGGCAGCTCTCGCAGGCCA
GGCGGGCGGCCAGGATCATGTCCACCAACATGCCAAGTGGTGGCGTCCCTGTCATCCTGGGCT
GGCGGGCTGCATCGCGGCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCCTGAGGCAGAGTTCAGGCTTACCGAATGCAGGCC
TATTCACCATCTGGACTTCCAGCATGCTGCAGGAGTGCAGGCAGTGCAGCCCTGATGATCGTAGGCATCGTCC
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCATGTCTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACATTCTGGATGTCACAGCTAACATGTACACCGCATGGTGG
GATGGTGCAGACTGTTCAGACCAAGGTACACATTGGTGCCTCTGTTGGCTGGCTGGCTGGAGGCC
TCACACTAATTGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCCTCAGGCCACAGTGGCTACAAGCCTGGAGGCTTCAGGCCAGCAGTGGCTT
TGGGTCCAACACCAAAACAAGAAGATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCAGCAGTATGTGTAATGCTCTAAGACCTCTCAGCACGGCGGAAGAAACTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTCTTCTGCTTTGACTCACAGCTGGAAAGTTAGAAAAGCCT
CGATTTCATCTTGGAGAGGCCAATGGCTTAGCCTCAGTCTCTGCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTCATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTACCCAAAGAAAATTGGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGCTCTCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAGATTTGTAATCTCCAGCCCATGATCTCGGTTTCTT
ACACTGTGATCTAAAGTTACCAACCAAGTCATTTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTG
CGCGGGTCAGAAATTGCCCTAGATGAATGAGAAAATTATTTTTAAATTAGCTAAATAGCTAA
AATAAAATAATGTTTAGAAATGATACACTATCTCTGAAATAGCTCACCCTACATGTGGATAGAAG
GAAATGAAAAATAATTGCTTGACATTGCTATATGGTACTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCTTAGCACTTTGGGAGGCTGAGGAGAAGGATCACTGAGCCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGCTCTACAAAATACAGAGGAGAAAATCAGCCAGTC
TGGTGGCATACACCTGTAGTCCCAGCATTCCGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACGTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTC
AATAAAAATAATGGAACACAGCAAGTCCTAGGAAGTAGTTAAACTAATTCTTAA

FIGURE 118

MSTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKA
STGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 119

GGAAAAACTGTTCTCTCTGTGGCACAGAGAACCCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAAACATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTTG
GTGGTGTGGAATGGTGGCACAGTGGTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTATGTCCCTCTGGCTTCATGATGGCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATTTCAT
CATCACGGCATGGTGGTGCATCCCTGTGAGCTGGTTGCCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAACGTGAGCTGGAGAAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTGTTCTGCTGCCGTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACTTCCCATGCACAACCCAAAAGTTACACCCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTGATTGTGATGTTTTAACTTACTATAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAATGGACCCCAAAGAAACTTGTATTACTGTTCTTAACTGCT
AATCTTAATTACAGGAACTGTGCATCAGCTATTATGATTCTATAAGCTATTCAAGCAGAATGAGATA
TTAAACCCATGCTTGTATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTATCATTACTCAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAATAGAACTCAACTATTGCTTTCAGGGAA
ATCATGGTAGGGTTGAAGAAGGTTACTATTGTTAAAAACAGCTTAGGGATTAATGCTTCCA
TTTATAATGAAGATTAAAATGAAGGCTTAACTCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCTAACTTCTTATCCTCTCCAGGGCTTTTT
CTTGTGTATTAATTAAACATTAAAAGCAGATATTGTCAGGGCTTGCATTCAAACGTCT
TTCCAGGGCTACTCAGAAGAAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGT
AAAATTTGTTGTATTGAAAGAAGAATGATGCAATTGACAAGAAATCATATATGTATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGAGGTTCTCAATATAAATAAAGAGCAGAAAATA
TGTCTGGTTTCAATTGCTTACCAAAAAACACAACAAAAAAGTTGTCCTTGAGAAACTCACCT
GCTCCTATGTGGTACCTGAGTCAAAATTGTCATTGTTCTGTGAAAATAAATTCCCTTGTCA
CCATTCTGTTAGTTACTAAAATCTGAAACTGTATTGTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTCAGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLLGGVGMVGTAVTVMPQWRVSAFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGILMCAASVMSFLAFMAILGMKCTRGTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGGGTGAAAGGCGATTGATGCAGCCTCGGGCCCTGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCCTGCCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCCTCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCGCCGTGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGTACACCTGGATCCAGGTGCGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAACTAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGTCTAAGAGTTGTTCACTGGCTCACTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTGGTATTACATTCAATGGAGCTGAATGTTCAAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCG
CACTTCTCTGTGAAGGACTTGTGAAGGAATTGGTGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTCACTTCTCGC
ATTATTGAAGAACTACCAAAATGCTTAATTTCATTTGCTACCTCTTTTATTATGCC
TTGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTACACTGTTAAATCTAGCATTATTCACTT
CTTCAATCAAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACCTTCTCATAGTCACATT
CTCTCAACCTATAATTGGAAATTGTTGGTCTTTGTTCTCTTAGTATAGCATT
AAAAAAATATAAAAGCTACCAATCTTGACAATTGTAATGTTAAGAATTTTTATATCTGT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAA SPQR LRG LLLL LQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFGSLSRLKCRNACCQRWYFTFNGAECGPLPIEAIYLDQGSPEMNSTINIHTS
SVEGLCEGIGAGLVDV A I W V G T C S D Y P K G D A S T G W N S V S R I I I E E L P K

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGCTCTCCTGCCCTGGCTCCAACGCAGCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACTGCTGGGTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATAACATCATGTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTGATGTTGCTGCGAATGCGGTGTTGGGATTATTTGTTCTGGAG
TGTTCTGCGTGGTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGTGTCATGCAACTG
GCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTGAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTATAGCCCCACTGCTTACTGACAATG
CTTTCTGCGGAACGGGATGCCCTAAGGGCTGTAGGTGAGGCAAATGGTATATGTGA
ATCTCAGAAATTACAGGAGATACCCCTCAAGTATATCTGCTGGTGTCTAGGTTGTCCTCGCT
ATAACAGCCCTCAAAAACCTTAAGTATACTCAATTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTAAATGGAATAGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCTATTCTTAACAAATACCTTCAGACCTGTGACAATT
TACGGAACTTGGATCTGCTCTATAATCAGCTGCATTCTCTGGATCTGAACAGTTGGGCTTG
CGGAAGCTGCTGAGTTTACCTTACGGCTAACTCCCTGAGAACCATCCCTGTGCAATTCCA
AGACTGCCAACCTGGAACCTTGGACTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTCAGCTCAAC
CTGGCCCTTTCCAAGGTTGGTCAGCCTCAGAACCTTACTTGCAGTGGATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCTTACAAAGGCTGATTTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCCTGAATCTGCAGCGCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAAGCAGAAATATTGTCCTCTGTAACACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTCCTGCAAGGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGGCAAAGACTACAGAGAGGTTGATCTGCCAG
GGCTCTCCAAAGCCGACGTTAACGCCAACGCTCCAGGCCGAAGCATGAGAGCAAACCCCT
TGCCCCCGACGGTGGGAGCCACAGAGCCGGCCAGAGACCGATGCTGACGCCAGCACATCT
TTCCATAAAATCATCGGGGACGCGTGGCTTTCTGTCCTGCTGTCATCCCTGCTGTTAT
CTACGTGTCATGGAAGCGGTACCCCTGCGAGCATGAGCAGCTGAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCCAGCACCCAGGAATTATGTA
GATTATAACCCACCAACACGGAGACCGAGATGCTGCAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAAACCATGTAACAAAGAGCTCTAAAGCT
GGGAAATAAGTGGTGTCTTATTGAACTCTGGTACTATCAAGGGAACCGCGATGCCCCCCCCCT
TTCCCTCTCCCTCACTTTGGTGGCAAGATCCTCTGTCCTGTCCTGTTAGTCATTCAATAACT
GGTCATTTCTCATACATAATCAACCCATTGAAATTAAATACCAAAATCAATGTGAAGCTT
GAACCTCGGTTAAATATAATACCTATTGTATAAGACCCCTTACTGATTCCATTAAATGTGCA
GTTTAAGATAAAACTTCTTCATAGGAAAAAAAAAA

FIGURE 124

MGFMVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGMVYCESQKLOEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTR
PVTMLRNLDSYNQHLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHEHNQFSKLNLAFFPRLVSLONLYLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFCVPNLQRLNLDNSKLTFIGQEILDWSISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTCGCTACTGCTGAATGTCCGTCCGGAGGAGGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAACAGCTCTGCTC
GGTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTGCAATGGTCACATATGAACATCTCGAGAGGGTGTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTT
TTAGCCAATCCAATGACCTAGTGAAGGTTAGATGCAAATGGAAGGAAAAGGAAACTGGAAGG
AAAACCAATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCAATATAACAAAGAGCAGCACTGGTGAATATGGGAGATTAA
ACCACTTATGATACTGAAACACTACTTGGTATTGAATACACCAACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTGAGTCTATATAAGGCTTTACC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRGMVRTALGIIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLCVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRIIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSLRLTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGGCGGCAGGAGAGCGGCCGGCGTCAGCTCTCGAC
CCCCGTGTCGGCTAGTCCAGCGAGGCGACGGCGCGTGGGCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTACGGGGGCTCGGGGGCATCGGCGGCCGTGGCC
CGGGCCCTGGTCAGCAGGACTGAAGGTGGTGGCGTGCACGGCCACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAAGGCTACCCGGACTTGTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTAGACATC
TGCATCAACAATGCTGGCTGGCCGGCTGACACCCCTGCTCTCAGGCAGCACAGCAGTGGTGGAA
GGACATGTTCAATGTAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATGTCGGCCACCGAGTGTAA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCCAGGTGTGGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGTGCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCGAGGCTTATCTACGTCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACTAGTGACTGTGGAGCTCC
TCCTCCCTCCCCACCCCTCATGGCTGCCCTGCCTCTGGATTAGGTGTTGATTCGGAT
CACGGGATACCACCTCCTGTCACACCCGACCAGGGCTAGAAAATTGTTGAGATTTATA
TCATCTGTCAAATTGCTCAGTTGTAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGTGCCCTGGCACTTGGCTTGTCTCAGTG
TCCTCCCTTGACATGGGAAAGGAGTTGTGGCAAATCCCACCTCTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCTCACCTTATCTGTGTTATCCAGGGCTCC
AGACTTCCTCCTGCCTGCCCACTGCACCCCTCTCCCCCTTATCTATCTCCTCTGGCTCCCC
AGCCCAAGTCTGGCTTCTGTCCCCCTCTGGGTCACTCCCTCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTGATGGTCAATTAAAAAGAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRILALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGTCTTCCTAGCCTCCTGCCGGTGGCTACACCAT
CATGTCCTCCCACCCCTCTTGACTGGGCCGTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCTCCGAGGCAGTCTGCTCAGAGGCCTCGGCCAGAATTCCAGTCTGGTTCATGC
CAGCCTGAAAGGCCATGGAACTTGGTGAATACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCGCTGTCATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTGAAATGAGGAACTGAGAAAATTAATTCTCATGTATTTCTCATTATTTA
TTATTTTAACTGATAGTTGTACATATTGGGGTACATGTGATATTGGATACATGTATA
TATATAATGATCAATCAGGGTAACTGGGATATCCATCACATCAAACATTATTTTATTCTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTATTTGTAGTAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGCCACCGTGCCCTGGCTAAACATTTCTTTCTTGTTGGAACTTGAATTAT
ACAATGAATTTGTTACTGTCATCTCCCTGCTGTGCTATGGAAACTGGAACTTCCCTCT
ATCTAACTGTTATTTGTACCAGTTACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGCTTCTGTGCTGGCTATTCAACATAATGACTCCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTGCTTCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSAREHLPSPRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMMPFKRVFCQDGTVRSFCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCATCGGGACTAAAATTGGGCTGTCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
TCCATCTGTGTGCGCTGCATGGGTTTCACTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACCTCTCACCTCAGAACACAAATAATAATGCTGGATTCCCTCAGAT
TTGAAAAAAACTTGCTGAAAGTAGAAAGATAACCTACACACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGAGAATTACATTAGATGACAACCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCCGACTGCTTCTGCCCCGTAATCACCTAGCACAAATTCCCTGGGTTT
GCCAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTTCATACCCATCTCTTCAAG
GTCTCACTAGTCTAAACGCTGGTCTAGATGGAAACCTGTTGAACAATCATGGTTAGGTGACAAAGTT
TTCTTCAACCTAGTTAATTGACAGAGCTGCCCTGGTGCAGAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCAACATCACTGGGTGCCCAAATGTTTTT
CTTATCTAAGGCAGCTCTACGACTGGATATGTCAATAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGATTCTCGAACAAATCCCTGGTATTGCGGGTCAAGATGAAATG
GGTACGTGACTGGTACAATCAGCTGGTGAAGGTCAACGTGCGTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGGGATGGCTATTAGGATCTCAATGCAGAACACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCAACTGCAATACCCAAACAGTGATCCCTGCCAAGGACAGTGGCCAGCTCAGTGAC
CAAACAGCCAGATATTAGAACCCCCAAGCTCAAGGATCAACAAACACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTCTGTCACCTCTGATACCATTCAATCTTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCAATGCCCGCATTTGGATCTATAACAGAAACAAATTG
AACAGGGGAAACGCAACTACCTACCTATTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTT
CCATGGGAAACCAACCTACAACCAACCCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCAATTAC
TTTGGCTGCCATCATTGGTGGGCTGTGGCCCTGGTTACCTTGCCTTCTGTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTCTCAAGGAACGTGCAATAGCAAAGGGAGGAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGAAATCAGGGAAACTCTTCAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGAGGAGTTGTAATACACACCATAATTCCCTCTAATGGAATGAATCTGT
ACAAAAAAACAATCACAGTGAAACGAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTGGGTTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAFIFYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLOENNIRTITYDLSKIPYL
EELHLDNSVSAV рееегаfrdsnylrllflsrnhsnstipwgлpтиеelrlddnristisspsl
QGLTSKLRLVLDGNLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPGQIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGIMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTQDQQTTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKGHSFAFGSITET
IVTGERSEYLVTALLEPDSPYKVCVPMETSNLYLFDETPVCETETAPLRMYNPTTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTATCCCCCTGCAGCCACCCCTCCAGAGTCCTTGCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGCCATGGCCAGCCTGGGCTGCTGCTCTGCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCTGTCT
GCGCTGGAGAGAGGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGTCCGAGTGTGCTGGAAGAGCAGCTAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTACCCAACTACCTAACAGAGAGTTCCAGCTGACCCCTCCAGCCCG
GTTTGGAAAGCTCCACATGCCCTGGATCCACACTGATGCCCTCTGGTGTACCCACGTTGGGC
CCCAGGACTCATCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCCTGCTCCACCAACTGCTCTTCTCCTGGGCCAGAATGAGGGATGCACACAGG
GACCACCTCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATGGATACGCCCTACCCCTACCGGGACATCTCATGGAAAACATCATGTTCTG
TGGAATGGCGGCTCTCGACTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGAGTGAAGAGGCCAGAAAAACAATTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTCTGCCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGTGCACCATACCTGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCATTCAAAGAGACTGTGTTGAATAAGGGCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCCTAGATGGCTGCTCTCCACAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCTATACATCTGGCAGAATACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
CCACCCACCAAGCAGCCGCTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTAGTCTCATCCCTAGATCCTGGAGGGCACGGATCACATCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCAATCTTGAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAACTGAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLGLARAQGAERVTSSEQRPMASLGLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMGVVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLCRSLMTKPGCGYCLSHOLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGTGGCATAGGAGATCGCTTCAGATTGAGGGCGTGCAGTTGTCAGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTTCATGATATACCTCTGGATTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCAGTGATATCACCTCGAAAGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTAGACTGCCCTATCCTCTCCAAATGAAATCTCAGGT
CCACCTCTTACTTTATTAAAGGAATCGTGGGGCTGGACAGACTTCTAAAGTGGTCAACACAAGTGATC
TATGATGATGGTCTTCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATCCAACCATGAGTTGCCTGAT
GTTCTGAGTTCATGACAAGACTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACACGGCAAAAGTGGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAACTT
AAACACGGCAACACTGGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGACTTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGCCTATAGAAAATGCCATTAATAATTAT
ATGAACTACTATACATTATGATATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLLTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPVMVVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMMNLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCCTCCCCAGTTCCCTGTGGGCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGCACTAACGGCGGAGCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTGAAATCTTCTTCACGGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGTCTCAGATTCAGGCCTAAGATGAAAGCCTTAGTCTTGCCCTCAGC
CTTCTCTGCGTTTATCTCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAATCTAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTTAA
AAACTACCAGACCCCTGACCATTATACTCTCGGAAGATCAGCAGCCTGCCATTCCCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGGAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGAAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGATGACCCAA
ACCACCATCTTTACTGTACTAGTCTTGCTGGTACAGTGATCTTATTGCTTACTTG
CTTCCTGCTGATGATTGTCTTATGCTGCCATTGAGACATCTGCTTACTTG
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTATTGCTATTAA
ATGTATTATTTTACTTGGACATGAAACTTAAAAAAATTCAAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTGACTACAGTAAAAAAACCTTGTAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAACTAAGGACATATTACTCATGCTGATGCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCATAATTGTGATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTTCTTGCAACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGCTGCAGCAGGGCAGGCCTCCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACGGATCGCGCTGGCTGGCTCCGGACTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGCTTCCTCCTCGGGATTCTCCCGGCTCCGCTCGTCCCTCGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGCGCCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTGAGAGATGATTTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACCTTGTGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGATTTGATGACGGGGAG
CCTTCCTGGCTTGTCGACGTACAGGAAACCTCAATTCTCGTCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAGAAATAGTCTTTATGGAGATGAACCTGGGTTAAATTA
TTCCCAAAGCATTTGTGGAATATGATGAAACACCTCATTTCTGTCGATTACACAGAGGT
GGATAATAATGTCACGGGCAATTGGATAAAGATTAAAAAGAGGAGATTGGGACATATTAAATCC
TCCACTACCTGGGGCTGGACACATTGGGACACATTTCAGGGGCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCCTTACCCAATTGCTGGTTCTTGTGGTGACCATGGCATGTCGAAACAGGAAGTCACG
GGGCCCTCACCAGGGAGGTGAATACACCTCTGATTAACTAGTTCTGCGTTGAAAGGAAA
CCCGGTGAATCCGACATCCAAAGCACGTCCTAGACGGATGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAGACAGTGTAGGGAGCCTCTATTCCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCATTTGAATACAGTCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAGATCCTGGTTGAGCAGTTAAATGTCAGAAAGATTGCAATGG
GAACCTGGATCAGACTGTACTGGAGGAAAAGCATTCAAGACTGCTTCAACCTGGGCTCCAAGG
TTCTCAGGCACTGGATGCTCTGAAAGACGCTGAGCTTGTCCCTGAGTGACAAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCC
CTGTCATCTCTGGTTTCTGCTTTATTGGTATCCTGGTTCTTCGGGCTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCGTTCTGACTCTCTGTGGCTCTCGTGGCTGGCGGAGGCT
GCCTTCGTTACCAACTCTGGTTGAACACTGGTGTGTCAGCTGGCAGTGCCCTGGAC
AGGGGCGCTCAGGAAAGGACGTGGAGCAGCTTACCCAGGCTCTGGGTGTCGGACACAGGTG
TTCACATCTGTGCTGTCAGGTCAAGTCAGTTCTGGAAAGCTAGTTCTGCACTGTTAC
CAAGGTATTGTAAGAGCTGGCGGTCAAGAGGAACAAGCCCCCAGCTGAGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGCACTGAGGGAAAGAGACAATGCCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCAACTCATCCTGCCACCCCCAGAATGCA
GCCTCATCAGGTCCAGATTCTTCAAGGCGGACGTTCTGTGGAATTCTTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGGTGAGGCACTGAGAAGAGGGCGGATGGTCAC
ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCACTGAGCAGGACTGTTGGCCCC
ACCCCAACCTGCACAGGCCCTCATCCCCCTTGGCTTACGGCTCAGGGCAGGATGATCTGTG
CTGACCGAGACACTCACAGCTTGTCACTGGGCACAGGCTCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCATCTGGCTCATGCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTTACCAAAAGAATAAACGGAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIROAKAAGKRIVFYGDETWKLFPKHFVEYDGTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGTTATCGTACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACA
ACTTCCCTATAGAAAACAAC TGCCAGCACCTTAAGACCACTCACACCTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCA GCATT CATGACCAGGATCACAAAGTACTGGTCCCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA ACTACATACGCCAGAGATCTTCTTGCA TTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAAACTGATGAAGCT
GGCTGCCAAAAGGAATCAGCACGCCGGCTTCATCTTATAGGGCTCAGGTGGCTCCTGG
ACATGCTGGAGTCGGCGCTCACCCGGATGGTCATCTGCACCTCCTGCAATTGTAATGACCT
GTTGGGGTGACAGATAAAATTGAGAACAGGAAACACATTGAATTTCATTCAACCAGTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCA CGCATTTAGGAAACTGCCCATGAAACGCCCTCGCTA
ATTGAACTAATTGTATAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLEYNFPPIENNQHLTTHTFRVKNLNPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGCTCTCCTGCTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGAAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATTGCTGGCCTGCCGTGG
AAGGGAGGTCTGCTCTGGCGCTGCTGCTGCTCTCTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCTACAGACTGGGG
CACATCTTGAATTCTGGAAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAGACGACATTGACAAC TGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTCACCTGCTTCTCACCACAGCACCAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTGAGTGAACACCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTAGAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGGATGAGGAAG
CACCTGAGCTGGTGGCTGGCCACTGTCTGCATGCTCTTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAACATCAAGTGGAACCGGAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGCTGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACCGAGGCCAACTACTGGCAGTCCCCGATGGCATTCCA
CTACAACGGCTGCTTGAGGCTAATGTGACCAAGGAGGCATTGTACCGGCTGCATCAATGCCA
CCCAGGGCGAACCCAGGGGAGTCCAGAACAGCAGAACAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTGCTCCCTCAAGCATTGCGAGTTGGTGGAGAGGGCGCAGGACT
TCGGGTACCATGCACCAAGCCAGTGCTCCTGCCTCTGGCTTGTATCTGGCTATGGTAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCCAGCAGCAGCAGCAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCGCTCCACGTATGCCTGGTATGCGCTGCGTCTGATAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACGCGCCACCTGCTCACAGTACTTCCAACACTCTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKCEFWLERGAGLRTMHQPVLLCLLALIWLKV

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGUCCACCATGTCGGTCCAGAGTCATTC
CCTGATGATTTAGACTCAAAGAAAACTGTTCAGAAGCTCTCTCTCTGGCCTCCTCT
CTGTCTTCTTCCTCTTCTTATTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCAAGTGGGTAGGTCACTGAGTCTTAGTTTTATTTTGAATTT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCTGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGGCCAACCTGCTTATCCCTGACCGTCAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCAACACCCACCCCTGGCTCTCCTGTTTACTCCTCTTCAATTCTA
ACAAAAGCTACAGCTCCAGGAGGCCAGCGCCGGCTGTGACCCAAGCCGAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTGGATTCTGGTGTAGTGCCTCCGATTCAAGCTTCCCAAACCTGGAGGAA
GCCAACAGAAATCTCTACATAATAGAGAATTAAAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGCAATAACAGAAAAGGAAAAATTGAGAAAGAACATCTA
TAAGAACGCTCCCCACTTGATAATAAGTGAATGTGGAGATGTTGATTCAACCAAGAACATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAATTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTAACCGCTGAAGACATTGTCCTAAATCGCTGCCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAAGATGAAACAGTATCTAACACA
TTAACCTTGACAATGGCTTGAAAGGAGAACTAAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTCCCAAATTCTATGCGCTACTGAAAAGTATTGATTCAAGAAAGCAGAACATGAGAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAACATGAGAA
CCAGAAGAAGGTGTTCTACCTGAAAACCTGGATGAAATGATTGCTCTCAGACCAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCATCAGAGAACAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACAGATGGAAAAGGAATATGGAAGCTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAACCTATTGGAAGC
CATCAGAAAAAATTGAATGGTGAAGAAACATGACAAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTAGCAGCCTGTAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTTCAATTCTGTGATTAAAATTGGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTTACAAGTGGTTAAAACATAGCTTCTTCCC
GTAAAAAACTATCTGAAAGTAAAGTTGTATGTAAAGCTGAAAAAAAAAAAAAAAAAAAA

1234567890

FIGURE 150

MGFLGTGTWILVLVLPIQAFFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNPKPTSWTENQAGKIKEVKTPMAAIQDGLAKGENDETVSNTLTITNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNLEKNATDNISKLFPAPEKSHHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKNGKEDYDLSKMRDFINKQADAYVEKGILD
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCCTGTGGAGCTCAA
GATGGTCCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCAATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCACTCTGGGTGTCAGGGTGGAAAG
CCAGTGCCTGTATGTGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTCCAAGGAATCCAAGAGCTTCACCTCTACCGGGGGACATGGGGCTCACC
TCCAGCTCGAGTCGGCTGCCAACCGGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCACTCACCCAGCTTCCCAGAGAATGGTGGCTGGAATGCCCATCACAGACTTCTACTTCC
AGCAGTGTGACT**AGGCAACGGCCCCAGAACACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGT**
GAGTGGAGGAGACCCATGGGGACAATCACTCTCTGCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCAACTTGTCTCTGGTCTCCAGTTGGATAAAATTCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGCCTACTGCTGTGGAACCTTGTAAAAACATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCCGCAAGGCCACCCATCCCCAGTTGAGCCTTATA
GGGTCACTGCTCTCCACATGAAGTCCTGTCACTCACCACGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTGGCCCAGCCCCACCCCTCCCTTAATCTGCCACTGTCA
TGCTACCTTCTATCTCTCCCTCATCATCTTGTGGCATGAGGAGGTGGTGTAGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTATTTAATTAAAAGATAACCTATTATATT
TCTTTATAGAAAAAAAGTCTGGAAGAGTTACTTCATGTAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTTCTTTAATTCTGTTAATTATCTGTATTTCTAATTTCTACAATGAAGATGA
ATTCCCTGTATAAAAATAAGAAAAGAAAATTATCTTGAGGTTAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTCCACTTCCCCAGAGTAAATTCAATTGATCAGGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGTAGCAGGAAACAGATCTCAGCAAAGCAGTCTGAGGAGGGCTGTGCTGAGTTGT
GTGGCTGGAATCTCTGGTAAGGAACCTAAAGAACAAATCATCTGTTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATGGATCCAGTCTAAGAAGGCTGCTACTGTTGA
ATTGTGCCCCCTCAAATTACATCCTCTTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGAGAGCACAGAGACAGAGGAGACGGGGAAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAAGGCCACAAGCTAAGAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGAAGAGGCAAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAAGAATAAATTCTGGCTGTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTGT
CTCCCTCGTGTACATTCTGTGTGTCCTCTCCACAATGTACCAAAGTTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTGGCATGCCACTTCAAGATTAGGTATAAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTGCCCCACCCAGGCCCCAATCTATCTGGCTCATCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAAGAGACTTACGGTGGAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAACAGAGACTCTGTGAGATAATCGATGTTGTGTTT
AAGTTGCTCAGTTGGTCTAACCTGTTATGCAAGCAATAGATAAATAATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEETSVVNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCCCTGGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCACCCTCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTGGCTGATAA
CAACACAGACGTTCGTCTCATTGGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCCTGAAGAAGTGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTCCCTGCTAGAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAATGAACCCCTGCCTAGTTACAAAGGAAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTCAATTGTAACGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGCTTTCCATAAAAAAGATTACTTCCAT
TCCTTAGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTTATATTATAAA
TGTATTATTATTATTATAAGACTGCATTTATTATCATTATTAAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAAGGCTAATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKS VSSFLMGT LATSCLLL ALLVQGGAA APISSH CRLDKS NFQQPY ITNRTF MLAKEASL
ADNN TDVRLIGE KLFHG VSM SERCYLMQ VLNFT LEEV LFPQS DRF QPYM QEV VPFL ARLSNRLS
TCHIEG DDLHI QRNV QKLKD TVKKLGE SGE IKAIG ELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGTGGAGAAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCGTAGGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACACGCCGT
TGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACCAGACTGTCTTACAGGCAGGCATGCCATGGCAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTGTGTGTGCAGGGCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCTGAAGGGCAGGATGCCAGATGCTTGGCCCTGTGAAGTGCTGTGGAGCAG
CAGGATCCCGGACAGGATGGGGGCTTGGGAAAACCTGCACCTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCCCTGTCTTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCTGTCTTCCCTTTCCATCCCTGCTACCTG
GCCCACAGCACAGGCACTTCTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTTGTTACTCATCACTCAAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGGAGGAAGCTGTTATTGAATGTATAGAGATTATCCAAATAAATAT
CTTATTAATAAATGAAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCP SKGQDTSEELL RWSTVPV PPLEPA
RPNRH PESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHC VSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVR PRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATGTCGCTCGTGTGCTAAGCCTGGCCGCGTGTGCAGGAGCGCCGTACCCGAGAGCC**
GACCGTTCAATGTGGCTCTGAAACTGGGCATCTCCAGAGTGGATGCTAACACATGATCTAATCC
CCGGAGACTGAGGGACCTCCGAGTAGAACCTGTTACAACTAGTGTGCAACAGGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGCAGATGCCAGCATTCCGCTTGTGAAGGCCACCAA
GATTGTGTGACGGCAAAAGCAACTTCCAGTCCTACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGGTAAATGGACATTTCCTACATGGCTTCCCTGTA
GAGCTGAACACAGTCTATTCAATTGGGCCATAATATTCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCGGCTGCCAGCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGATCCGAACATCACTGCTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCACCCCCGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATCGGGTTTCTCAGGTGTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGAATGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTCCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCTCT
GGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCCTCCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGGCACGAAAGGATCAAGAAGACTTCC
TCTACCACCAACTACTGCCCTCATTAAGGTTCTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGAGAAGTGGTCTGCTGGTGGCCA
AGTGGCAGAAAAGAAAATAGCAGAGATGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGCTTCCCTTCCATGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGTCACTTAGAGAGATTGATA
AAAGACGATTACAATGCTCTCAGTGTGCTGCCCAAGTACCTCATGAAGGATGCCACTGCTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCCTG**TAG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNVS WV
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTVGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDN NK
SKPGGWPLLLLSSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGAACATGACAGTGAAGACCCATGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGCTTGCCTTCTGAGTGAGGCGGAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTCCAAAAGCCTGAGAGTTGCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTGGTCCGGAGGAAGCACCAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPAMVKYLLLSILGLAFLSEAAARKIPKVGHFFQKPESCPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGTCAGGACTCCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTGGAGGCTGAAGAGGGATTG
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTGTCCT
TGGCACTGGGCGGAAGCCCAGTGGCTTCTGGAGAGGCTTGTTGGGCCCTCAGGACGCTACC
CACTGCTCTCCGGGCTCTCCGCGCTCTGGGACAGTGACATACTTGCCCTGCCCTGGGACAT
CGTGCCTGCTCCCCGGCCCTGCTGGCGCTACGCACCTGCAGACAGCTGTCAGGTGCTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGCTGCGTGCTGGCTGCACTTGGCGTGCATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTGACTCAGGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCAAGTGTGCTCTCCAGGCCACCTACTTGCCCGCTGGTCCCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTTGTGCACTGGTCACTGTGGCTCTGTGGTATATGAC
TGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTCTTACTCAGCCCAGGTACGAGAA
GGAACATCAACCACACAGCAGCTGCCCTGCCCTGGCTCACGTGTACGAGATGGTACA
ACGTGCATCTGGTCTGAATGTCTCTGAGGAGCAGCACTTCCGCTCTCCCTGTACTGGATCAG
GTCCAGGGCCCCCCTAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATCATTACCTGAA
CCACACAGACCTGGTCCCTGCCCTGTATTCAAGGTGTCAGCTGGCCTCTGGAACCTGACTCCGTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGAGAGCTGGCTGCTGGACGCACTGGCCTGCTGCCCGCAGAACGGCACT
GTGCTGGCGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCACCGCTTCTGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTCCATTGCTGAAAGGCCACCCCTAACCTCTGTGTTAGGTG
AACAGCTGGAGAAGCTGCAAGCTGCAAGGAGTGCTTGTGGCTGACTCCCTGGGCCCTCTCAAAGA
CGATGTGCTACTGTGGAGACACGAGGCCAGGACAACAGATCCCTGTGCTTGGAACCCA
GTGGCTGACTTCACTACCCAGCAAAGCCTCCACGAGGGCAGTCGCCCTGGAGAGTACTTACTA
CAAGACCTGCACTGAGGCCAGTGTCTGCACTATGGGACGATGACTTGGAGCGCTATGGCCTG
CCCCATGGACAATAACATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCTGCCACTCTTGCCG
CTGCGCTTCCCTCATCCTCTCAAAAAGGATCACGCAAAAGGGTGGCTGAGGCTCTGAAA
CAGGACGTCGCTCGGGGGCGGCCAGGGGCCGGCTCTGCTCCTACTCAGCCGATGA
CTCGGGTTCTGAGCGCCTGGTGGGCCCTGGCGTGGCCCTGTGCTGCCAGCTGCCGCTGCCGTTGG
CCGTAGACCTGTGGAGGCCGTCGTGAACTGAGCGCGCAGGGGCCGTGGCTTACCGCGCAG
CGGCCAGACCTGCAAGAGGGCGGCCGTGGTGTCTGCTCTCTCCCGTGCCTGGCTCTGAAA
GTGCAGCGAGTGGCTACAGGATGGGTGCTCCGGGCCGGCGCACGGCCCCACGACGCCCTCC
GCGCTCGCTCAGCTGCGTGTGCCCCGACTTCTTGCAAGGGCCGGCGCAGCTACGTGGGG
GCCCTGCTTCGACAGGCTGCTCCACCCGGACGCCGTACCGCCCTTTCCGCACCGTGCCGTCTT
CACACTGCCCTCCAACTGCAAGACTTCTGGGGGCCCTGCACTGAGCAGCCTCGGCCCGCTGG
GGCGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGACTCCCGCGCCGGACGCCGGGGTGGGACCGAGGGCGGGACCTGGGGGGGG
CGGGACTTAAAAAGGCAAGCAGCTGTTTCTAAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVPQDATHCSPGLSCLRWDSDILCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAHVHLAVHGHWEEPEDEEKFGGAADSGVEEPRNASLQAQVVLFSQAYPTARCVLLEVQVPAALVQFGQSVGSVYDCFEAALGSEVRISYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQVNSSEKLQLQECLWADSLGPLKDDVLLLETGPQDNRSCLALEPSGCTSLPSKASTRAARLGEYLLQDLQSQCQLQWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLILLLKKDHAKGWLRLKQDVRSGAAARGRAALLLSSADDGFERLVLGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVLLFSPGAVALCSEWLQDGVSQPGAHGPDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVVFPLPSQLPDFLQALQQPRAPRSGRQERAEQVSRALQPALDSYFHPPGT PAPGRGVGVPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394.

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTGGATCTGCTCCAGCACGTGAAATTCCAGTCCAGCAACTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTCAGGGATCACCCGGAAGTCCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCAGGGAGGCGGTAGCCACCAAGATGACTGACAGGGTCAAGCTCTGAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGGAGATCGATTAGATGATTGTCATCCTA
CCCCCACGCAATCCGTGCAAGCGATGGCCACCGGCTAACCTGGAAAGACATCTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACTTGGAGGGAAAGCAGAGAGA
ATATGAGTTCTCGGCCCTGACCCCTGACACAGACTTCTTGGCACCATCATGATTGCTTCCCA
CCTGGCCAAGGAGAGTGCCTTACATGTGGAGGTGAAGACACTGCCAGACGGGACATGGACC
TACTCCTTCCGGAGCTTCTGTGATGGCTTCCATGGCTTCCCTGTCAGTACTCTGACTGAG
CTACAGATATGTCACCAAGGCCCTGCACCTCCAACTCCCTGAAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGGCTTCACTCAGGAGCACGTCTGACCCCTGCTTGGACCTCAGCGGCCAGC
AGTCTGGCCAGGCTGTCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGGCCAGGAGC
TCCACAGCGGCTAGCCCTGTCAGTACTTACCTAGGGCAGCCAGACATCTCCATCCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCATCCTATGCACTCAGGTGACCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTAAGGTCCAGCCTCTCTATGCCCTCAAGCCACTCCGGACAGCTGCCCTCC
CCTATGGGTATGCATGGAAAGGTTCTGCAAAGACTCCCCACTGGGACACTTCTAGTCTAA
CACCTTAGGCCTAAAGGTCAAGTTCAGAAAGAGGCCACCAGCTGGAAAGCTGCATGTTAGGTGGCCT
TTCTCTGAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAAAATCATTGCAACCAGC
CCCTGGGATTTGCACAGACAGAACATCTGACCCAAATGTGTCACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCAGTCCAGATCGAGGGCCACCCATGTC
CCTCCCTTGCAACCTCTCCGGTCCATGTTCCCCCTCGGACCAAGGTCAAGTCCCTGGG
TGCTGGAGTCCCTGTGTGTCCTAGGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAACTGGATTCTTTCAGAGGCTGGCCCTGACTGTGCACTGGAGTCCCTG
AGGGAATGGGAAGGCTTGGTGTCTCCCTGTCCTACCCAGTGTACATCTTGGCTGTC
ATCCCATGCCCTGCCATGCCACACACTCTGCATCTGGCTCAGACGGTGCCTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGTGCCTCTCAGGGAAACAGCATGATA
AGGACTGCAGCGGGGAGCTGGGAGCAGCTTGTGAGACAAGCGCTGCTGAGGCCCTG
CAAGGCAGAAATGACAGTCAAGGAGGAATGCAGGGAAACTCCCAGGTCCAGGCCACCTC
CTAACACCATGGATTCAAAGTGTCTAGGGAAATTGCTCTCTGTGTCATTGTTCAAGGTGG
ACAATCTAGCTGACAGAGCATGAGGCCCTGCCCTGCTCTGTGTCATTGTTCAAGGTGG
GCCCTGGAAAAGAACCAAGGCCAGGGCTGGAAAAGAACAGAAGGGCTGGCAGAAC
ACTTCTGCCAAGGCCAGGGCAGCAGGACGGCAGGACTCTAGGGAGGGTGTGGCTGAGCTCA
TTCCCAAGCCAGGGCAACTGCTGACGTTGACGATTTCAGCTTCATTCTCTGATAGAAC
GAAATGCAGGTCCACAGGGAGGAGACACACAAGGCTTTCTGAGGCAGGAGTTCA
ATCTGAGAATGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAACAC
TGTACTGATGTACAACATTGCAAGCTGCTGGTTCAAGCCATCTGGCTCAAATTCCAGC
CTCACCAACTACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCTGAGAACCTCGG
ATCTGTAATGTGGGATCATACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGT
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCTGAAATAACGGTAGCTATT
AAAAAAA

43063687 - 032022

FIGURE 164

MRTLLTILTVGSIAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILOPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLLSPKHLRPKGQLQKEPPAGSCMLGGSLQEVTSAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLGQLPLSSVQIEGHPMMSLPLQPPSGPC
SPSDQGPSPWGLESVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCCGGGTGGCCACAACATGG
CTGGCGCCGGGCTGCTCTTCTGGCTTGTGCTGGGGCGCTCTGGTGGTCCCAGGCCAG
TCGGATCTCAGCACGGACGGCTTCTGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTTGTAAAGACTTCACGGCCCTGATTGTCGTTGTGAATTAA
AAAAAGGTGACGATGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGGTGAACACAGTTGGATATTCACAAAGATTGATCAAGGTACTTCATAAAACACGGA
AGAAGAGCTACATAATTCCAGCAGATGAGACAGACTTTGTCGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCTGAGGAGTCTCGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCGAGGCATTAGCAGACTTGTCGCTTGAAGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAATCTAAGTTGTTACAAAGATTGTTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTGAGCAAACAAAAATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVC GDEEC SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAEEVSQHREKSPPEESRGRELDPVPEPEAFRADSEDGEGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACCGCGCACCTCGAAGTCTGAACTCCAGCCCCGACATCCACCGCGGGCACAGG
CGCGGCAGGCAGGTCCCGCGAAGGCATGCGCGCAGGGGGTGGCGAGCTGGGCTCGGGC
GGCAGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGTGCATATTCAAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCGCTCGCTCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGGCCGCTGG
TCTCGCTGCTGCTGGCGCCGCTGCTGCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GGCCAAAAGGTGTGTTTGCTGACTTCAGCATCCCCTGCTACAAAATGGCCTACTTCCATGAACT
GTCCAGCGAGTGAGCTTCAGGAGGCACGCCCTGGCTTGAGAGTGAGGGAGGAGTCCCTCA
GCCTGAGAATGAGCAGAACAGAACAGTAAAGAGAGCATGTTGCAAAACCTGACAAAACCGGG
ACAGGAAATTCTGATGGTATTCTGGATAGGGCTTGGAGGAATGGAGATGGGAAACATCTGG
TGCCTGCCAGACTCTACCACTGGCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTCTCGGGAAGTGAAGGAAAGTGTGTTGTGATGTATCACAAACCAACTGCCAATCTGGC
CTTGGGGTCCCCACCTTACCACTGGATGATGACAGGTGAACTGAAGCACAATTATTTG
CAAGTATGAACCAGAGATTAACTCAACAGCCCCCTGAGAAAAGCCTTACAAATCAACCA
GAGACACCCATCAGAATGTGTTACTGAAGCAGGTATAATTCCAACTCTAATTATGTTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTGTCTTGGAAACCTGTTGCTTCCAGATGCT
GCATAAAAGTAAGGAAGAACAAAATGTCACACTGTGGATTTCAGAAGTAA
CCAGAAAAGAAAGTGGCATGAAAGTATAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGAATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGAAATATTAAAGTAATTGTTATGCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAA
ACTTCAAACTCAAGCAAATGAAATGGACAATGCAAGATAAAGTTGTTACACACGTCGGGAGTA
TGTGTTAGAGCAATTCTTTTATTCTTACCTTCAAACTGTTGATGAAAGTATTACCTTGCATAAGTGT
TGTATATTGTTGAAATTACAGTGTGAAAAGTATTACCTTGCATAAGTGTGATAAAA
ATGAACCTGTTCAATATTATTATTGTCATCTCATTTCAATACATGCTTTGATTAAAG
AAACTTATTACTGTTGTCACGAAATTGACAATGCAAGATAAAGTTGTTACACACACAAATATGAC
TTTCTCGAAATAATTCTCATTTCAAGCTCTGCTTTGGTCATGCTAGGAATCTTCA
AATAAGAAGCTATTCTTCAATTAGTGTGATATAAACCTCCTCAAACATTTACTTAGAGGCAAGGAT
TGTCTAATTCTCATGTGCAAGACATGTCCTTAAATTATTGCTTAAATTAAACAGATT
TTGTAATAATGTAACCTTGTAAATAGGTGCAAAACACTAATGCACTCAATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGTCCTCACACGTCCTATATAATGAGAAGCAGCTCTGA
GGGTTCTGAAATCAATGTGGCCCTCTTGCCTAAACAAAGATGGTTGTTGGGGTTGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAACAAAGAGGTCACTGGTTGAGGACAGGTGAAATGTCACATCAGTGT
ACAAGCAGCACACAGACATTAGGAAGGAAAGGAAACTACGAAATCGTGTGAAAATGGTTGG
AACCCATCAGTGATCGCATATTCAATTGAGGTTGAGGTTGAGGATAGAAAATGGTGGCTCCTT
CTGCTTATCTCTAGTTCTCAATGCTTACGCCCTGTTCTCAAGAGAAAAGTTGTAACCT
CTGGTCTTCATATGTCCTGCTCCTTTAACCAATAAGAGTTCTGGGGAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVSLLLGAALLCGHGAFCRRVSGQKVCADFKHPCYKMAFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSQACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217